

The background features a gradient from light blue on the left to light yellow on the right. Overlaid on this are several large, semi-transparent shapes: two vertical yellow bars, a large red circular shape, and a horizontal red bar. The text is centered over these shapes.

VR-6 Why is my Reserve Distribution Too Narrow?

Chris Gross
Christopher Gross Consulting
(651)293-8008

Antitrust Notice

- **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.**

Presentation

- Correlation
 - Between development periods
 - Between accident periods
 - Between methods
- Effect of skewed distributions
- Tail Factors
- The role of judgment

Basic Data

- A typical triangle of losses will serve as an example

	Age1	Age2	Age3	Age4	Age5	Age6	Age7	Age8	Age9	Age10
2000	2,800,405	3,768,316	4,045,263	4,236,909	4,289,114	4,321,256	4,336,731	4,344,159	4,348,544	4,348,968
2001	2,608,018	3,573,193	3,991,428	4,194,788	4,256,829	4,286,767	4,313,103	4,317,923	4,318,432	
2002	2,551,421	3,596,466	3,973,491	4,158,497	4,212,343	4,240,059	4,254,540	4,260,569		
2003	2,200,963	2,993,323	3,260,122	3,388,538	3,430,394	3,446,614	3,454,098			
2004	2,220,241	3,076,331	3,314,883	3,437,104	3,481,601	3,498,406				
2005	2,518,198	3,482,141	3,732,331	3,867,741	3,910,920					
2006	2,624,262	3,596,594	3,877,256	3,993,855						
2007	2,823,468	3,872,470	4,126,395							
2008	2,443,839	3,176,659								
2009	1,954,178									

Wtd Average	1.366142	1.084493	1.041331	1.012783	1.006244	1.003914	1.001416	1.000565	1.000098	
Age to Ult	1.581721	1.157801	1.067596	1.025223	1.012283	1.006002	1.00208	1.000663	1.000098	

- For this illustration we will assume development is finished at ten years, and certain in the tenth year

Variability of Development Factors

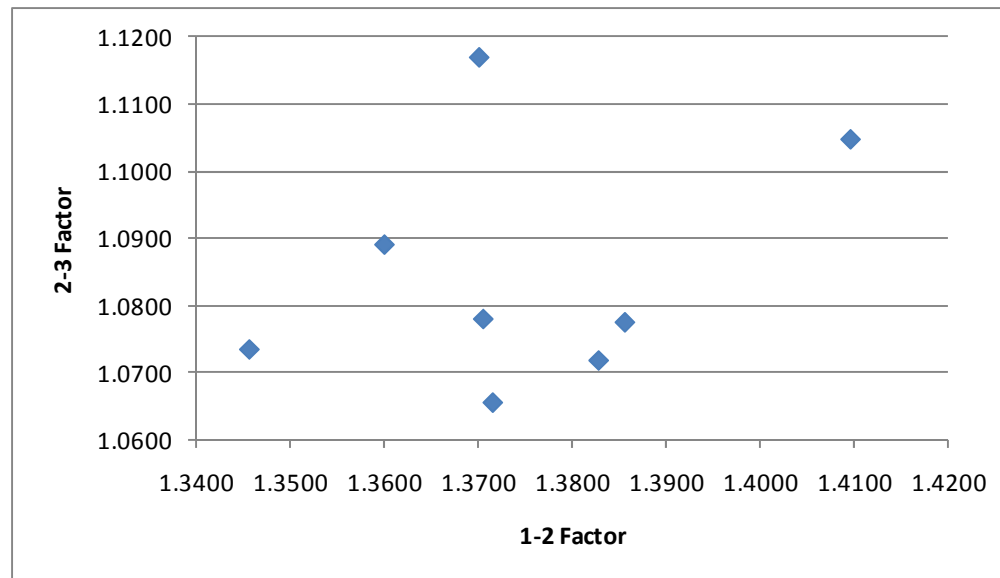
	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10
2000	1.3456	1.0735	1.0474	1.0123	1.0075	1.0036	1.0017	1.0010	1.0001
2001	1.3701	1.1170	1.0509	1.0148	1.0070	1.0061	1.0011	1.0001	
2002	1.4096	1.1048	1.0466	1.0129	1.0066	1.0034	1.0014		
2003	1.3600	1.0891	1.0394	1.0124	1.0047	1.0022			
2004	1.3856	1.0775	1.0369	1.0129	1.0048				
2005	1.3828	1.0718	1.0363	1.0112					
2006	1.3705	1.0780	1.0301						
2007	1.3715	1.0656							
2008	1.2999								

SD of Age-Age Factors 0.030533 0.017823 0.007437 0.001191 0.001279 0.001667 0.000298 0.00063
 SD of Age-Ultimate Factors 0.045487 0.020948 0.008078 0.002561 0.002231 0.001812 0.000698 0.00063
 (assuming independence)

- Combination of variability factors without making a distributional assumption, but assuming independence

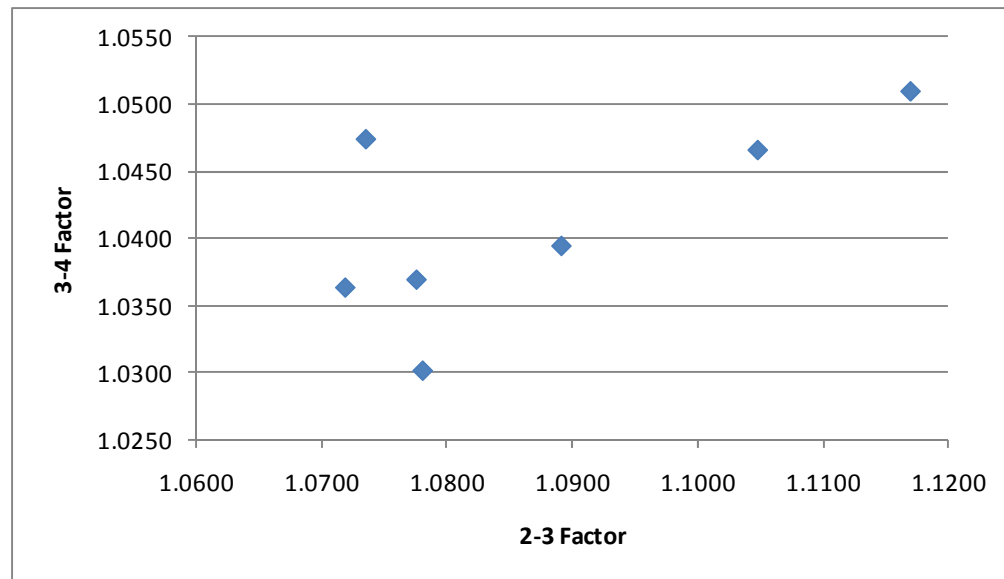
Independent?

- Scatter plots of paired historical development factors can illustrate lack of independence.
- Here the second development factor is shown against the first one.



Independent?

- ... and here the third development factor against the second



Correlation matrix

- A correlation matrix can be created from looking at such pairs:

Correlation Coefficient	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9
1-2	1.000	0.305	(0.056)	0.038	(0.173)	0.047	(0.375)	(1.000)
2-3	0.305	1.000	0.668	0.836	0.277	0.655	(0.968)	(1.000)
3-4	(0.056)	0.668	1.000	0.705	0.924	0.902	(0.768)	(1.000)
4-5	0.038	0.836	0.705	1.000	0.319	0.931	(0.963)	(1.000)
5-6	(0.173)	0.277	0.924	0.319	1.000	0.642	0.501	1.000
6-7	0.047	0.655	0.902	0.931	0.642	1.000	(0.840)	(1.000)
7-8	(0.375)	(0.968)	(0.768)	(0.963)	0.501	(0.840)	1.000	1.000
8-9	(1.000)	(1.000)	(1.000)	(1.000)	1.000	(1.000)	1.000	1.000

- Unfortunately, statistical significance is often lacking due to limited data (ex. 8-9 factor)

Lognormal Distribution Model

- Not always usable, but helpful to illustrate effect of correlation on reserve variability

Ln(Age to Age Factor)	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10	
2000	0.2969	0.0709	0.0463	0.0122	0.0075	0.0036	0.0017	0.0010	0.0001	
2001	0.3149	0.1107	0.0497	0.0147	0.0070	0.0061	0.0011	0.0001		
2002	0.3433	0.0997	0.0455	0.0129	0.0066	0.0034	0.0014			
2003	0.3075	0.0854	0.0386	0.0123	0.0047	0.0022				
2004	0.3261	0.0747	0.0362	0.0129	0.0048					
2005	0.3241	0.0694	0.0356	0.0111						
2006	0.3152	0.0751	0.0296							
2007	0.3159	0.0635								
2008	0.2623									
Average	0.3118	0.0812	0.0402	0.0127	0.0061	0.0038	0.0014	0.0006	0.0001	
Std Dev	0.0226	0.0163	0.0071	0.0012	0.0013	0.0017	0.0003	0.0006		
Sum of Ln(Factor)			0.4579	μ				$\text{Exp}(\mu + .5*\sigma^2)$	1.5814	
(Sum of SD(Ln(Factor)))^2)^.5			0.0289	σ				$\text{Exp}(\mu + .5*\sigma^2)*(\text{Exp}(\sigma^2)-1)^.5$	0.0457	
									Estimated 2008 AY SD (Assumes Independence)	89,277

Residual Terms

Difference of Ln from Average	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9
2000	(0.0149)	(0.0103)	0.0061	(0.0004)	0.0014	(0.0002)	0.0003	0.0004
2001	0.0031	0.0295	0.0095	0.0020	0.0009	0.0023	(0.0003)	(0.0004)
2002	0.0315	0.0185	0.0053	0.0002	0.0004	(0.0004)	0.0000	
2003	(0.0043)	0.0042	(0.0016)	(0.0004)	(0.0014)	(0.0017)		
2004	0.0143	(0.0065)	(0.0040)	0.0002	(0.0013)			
2005	0.0123	(0.0118)	(0.0046)	(0.0016)				
2006	0.0034	(0.0060)	(0.0106)					
2007	0.0041	(0.0177)						
2008	(0.0495)							

- These terms will be used to develop a covariance matrix

Testing the hypothesis that $\rho=0$

- t statistic is useful:

$$r * \sqrt{ (n-2) / (1-r^2) } \sim t(n-2)$$

where r is the sample correlation coefficient and n is the number of observed pairs.

Grouping by Lags

1 period apart

(0.0149)	(0.0103)	(0.0016)	(0.0004)	measured correlation
0.0031	0.0295	(0.0040)	0.0002	0.362136
0.0315	0.0185	(0.0046)	(0.0016)	
(0.0043)	0.0042	(0.0004)	0.0014	number of observations
0.0143	(0.0065)	0.0020	0.0009	35
0.0123	(0.0118)	0.0002	0.0004	
0.0034	(0.0060)	(0.0004)	(0.0014)	t-statistic
0.0041	(0.0177)	0.0002	(0.0013)	2.23
(0.0103)	0.0061	0.0014	(0.0002)	
0.0295	0.0095	0.0009	0.0023	probability under null hypothesis
0.0185	0.0053	0.0004	(0.0004)	0.032535
0.0042	(0.0016)	(0.0014)	(0.0017)	
(0.0065)	(0.0040)	(0.0002)	0.0003	
(0.0118)	(0.0046)	0.0023	(0.0003)	
(0.0060)	(0.0106)	(0.0004)	0.0000	
0.0061	(0.0004)	0.0003	0.0004	
0.0095	0.0020	(0.0003)	(0.0004)	
0.0053	0.0002			



Correlation by Lag

- Repeating with other lags, results are summarized below:

lag	measured correlation	probability under null	selected correlation
1	0.362	0.0325	0.3621
2	0.076	0.7071	0.0000
3	0.195	0.3972	0.0000
4	0.202	0.4896	0.0000
5	-0.061	0.8767	0.0000
6	-0.698	0.1903	0.0000

- May wish to add reasonable constraints such as that the correlation for a given lag is no greater than a smaller lag.

Correlation Matrix

- Build selected correlation matrix

Correlation Matrix	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9
1-2	1.0000	0.3621	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2-3	0.3621	1.0000	0.3621	0.0000	0.0000	0.0000	0.0000	0.0000
3-4	0.0000	0.3621	1.0000	0.3621	0.0000	0.0000	0.0000	0.0000
4-5	0.0000	0.0000	0.3621	1.0000	0.3621	0.0000	0.0000	0.0000
5-6	0.0000	0.0000	0.0000	0.3621	1.0000	0.3621	0.0000	0.0000
6-7	0.0000	0.0000	0.0000	0.0000	0.3621	1.0000	0.3621	0.0000
7-8	0.0000	0.0000	0.0000	0.0000	0.0000	0.3621	1.0000	0.3621
8-9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.3621	1.0000

Covariance Matrix

- Build the covariance matrix using $\sigma_{i,j} = \rho_{i,j}\sigma_i\sigma_j$

Covariance Matrix		Standard Deviation							
		0.02259	0.01633	0.00715	0.00118	0.00127	0.00166	0.00030	0.00063
Standard Deviation	0.02259	0.00051	0.000134	0	0	0	0	0	0
	0.01633	0.000134	0.000267	4.23E-05	0	0	0	0	0
	0.00715	0	4.23E-05	5.11E-05	3.04E-06	0	0	0	0
	0.00118	0	0	3.04E-06	1.38E-06	5.41E-07	0	0	0
	0.00127	0	0	0	5.41E-07	1.62E-06	7.64E-07	0	0
	0.00166	0	0	0	0	7.64E-07	2.76E-06	1.79E-07	0
	0.00030	0	0	0	0	0	1.79E-07	8.83E-08	6.782E-08
	0.00063	0	0	0	0	0	0	6.78E-08	3.97E-07

- Summing elements of sections of this matrix provides σ^2 for the age-to-ultimate lognormal.
- Highlighted area represents variability from age three.

Impact of Correlation

Year	Age	μ	σ	mean =	st dev =	Current Loss	Estimated	Standard
				$\exp(\mu+5\sigma^2)$	$\text{mean} * [\exp(\sigma^2) - 1]^{.5}$		Deviation of	Deviation
							Reserve and	Assuming
							Ultimate	Independence of
								Development
								Periods
2001	9	0.0001		1.0001	0.0000	4,318,432	-	
2002	8	0.0007	3.97E-07	1.0007	0.0006	4,260,569	2,686	2,686
2003	7	0.0021	6.21E-07	1.0021	0.0008	3,454,098	2,727	2,411
2004	6	0.0059	3.73E-06	1.0059	0.0019	3,498,406	6,800	6,335
2005	5	0.0120	6.88E-06	1.0121	0.0027	3,910,920	10,380	8,723
2006	4	0.0247	9.34E-06	1.0250	0.0031	3,993,855	12,513	10,226
2007	3	0.0649	6.65E-05	1.0671	0.0087	4,126,395	35,903	33,329
2008	2	0.1461	0.000418	1.1575	0.0237	3,176,659	75,152	66,183
2009	1	0.4579	0.001195	1.5817	0.0547	1,954,178	106,884	89,277
Combination of Accident Years (Assuming Independence)							136,698	117,026

Accident Year Correlation

- Residual terms from applying age to age factors to history to project historical losses one period forward and subtract actual loss

57,435	41,451	(24,451)	1,954	(5,361)	1,438	(1,286)	(1,931)
(10,270)	(116,324)	(38,390)	(8,421)	(3,358)	(9,558)	1,289	1,931
(110,862)	(73,148)	(20,777)	(690)	(1,414)	2,114	(3)	
13,505	(13,883)	6,328	1,458	5,199	6,006		
(43,166)	21,377	14,787	(562)	4,934			
(41,925)	44,028	18,851	6,261				
(11,479)	23,226	43,652					
(15,211)	73,273						
161,972							

Accident Year Correlation

- Correlation across accident years can be observed within a calendar period

57,435	41,451	(24,451)	1,954	(5,361)	1,438	(1,286)	(1,931)
(10,270)	(116,324)	(38,390)	(8,421)	(3,358)	(9,558)	1,289	1,931
(110,862)	(73,148)	(20,777)	(690)	(1,414)	2,114	(3)	
13,505	(13,883)	6,328	1,458	5,199	6,006		
(43,166)	21,377	14,787	(562)	4,934			
(41,925)	44,028	18,851	6,261				
(11,479)	23,226	43,652					
(15,211)	73,273						
161,972							

Accident Year Correlation

- Correlation across accident years can be observed within a calendar period

57,435	41,451	(24,451)	1,954	(5,361)	1,438	(1,286)	(1,931)
(10,270)	(116,324)	(38,390)	(8,421)	(3,358)	(9,558)	1,289	1,931
(110,862)	(73,148)	(20,777)	(690)	(1,414)	2,114	(3)	
13,505	(13,883)	6,328	1,458	5,199	6,006		
(43,166)	21,377	14,787	(562)	4,934			
(41,925)	44,028	18,851	6,261				
(11,479)	23,226	43,652					
(15,211)	73,273						
161,972							

Accident Year Correlation

- Correlation across accident years can be observed within a calendar period

57,435	41,451	(24,451)	1,954	(5,361)	1,438	(1,286)	(1,931)
(10,270)	(116,324)	(38,390)	(8,421)	(3,358)	(9,558)	1,289	1,931
(110,862)	(73,148)	(20,777)	(690)	(1,414)	2,114	(3)	
13,505	(13,883)	6,328	1,458	5,199	6,006		
(43,166)	21,377	14,787	(562)	4,934			
(41,925)	44,028	18,851	6,261				
(11,479)	23,226	43,652					
(15,211)	73,273						
161,972							



Grouping by Lags

Lag = 1

(10,270)	41,451	14,787	1,458	measured correlation 0.65144
(110,862)	(116,324)	18,851	(562)	
13,505	(73,148)	43,652	6,261	number of observations 35
(43,166)	(13,883)	(8,421)	(5,361)	
(41,925)	21,377	(690)	(3,358)	t-statistic 4.93
(11,479)	44,028	1,458	(1,414)	
(15,211)	23,226	(562)	5,199	probability under null hypothesis 2.25E-05
161,972	73,273	6,261	4,934	
(116,324)	(24,451)	(3,358)	1,438	
(73,148)	(38,390)	(1,414)	(9,558)	
(13,883)	(20,777)	5,199	2,114	
21,377	6,328	4,934	6,006	
44,028	14,787	(9,558)	(1,286)	
23,226	18,851	2,114	1,289	
73,273	43,652	6,006	(3)	
(38,390)	1,954	1,289	(1,931)	
(20,777)	(8,421)	(3)	1,931	
6,328	(690)			



AY Correlation by Lag

- Repeating for each lag and summarizing:

AY Lag	Measured Correlation	count	t	Probability Under Null	Selected Correlation
1	0.651	35	4.932	0.0000	0.651
2	0.579	27	3.551	0.0016	0.579
3	0.559	20	2.857	0.0105	0.559
4	0.426	14	1.630	0.1291	0.000
5	0.496	9	1.511	0.1744	0.000
6	0.189	5	0.333	0.7614	0.000

Correlation and Covariance Matrices

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	AY SD
2001	1.000	0.651	0.579	0.559	0.000	0.000	0.000	0.000	0.000	0.000	-
2002	0.651	1.000	0.651	0.579	0.559	0.000	0.000	0.000	0.000	0.000	-
2003	0.579	0.651	1.000	0.651	0.579	0.559	0.000	0.000	0.000	0.000	2,686
2004	0.559	0.579	0.651	1.000	0.651	0.579	0.559	0.000	0.000	0.000	2,727
2005	0.000	0.559	0.579	0.651	1.000	0.651	0.579	0.559	0.000	0.000	6,800
2006	0.000	0.000	0.559	0.579	0.651	1.000	0.651	0.579	0.559	0.000	10,380
2007	0.000	0.000	0.000	0.559	0.579	0.651	1.000	0.651	0.579	0.559	12,513
2008	0.000	0.000	0.000	0.000	0.559	0.579	0.651	1.000	0.651	0.579	35,903
2009	0.000	0.000	0.000	0.000	0.000	0.559	0.579	0.651	1.000	0.651	75,152
2010	0.000	0.000	0.000	0.000	0.000	0.000	0.559	0.579	0.651	1.000	106,884
AY SD	-	-	2,686	2,727	6,800	10,380	12,513	35,903	75,152	106,884	

Combined SD 205,845

Comparison of Results

- Standard Deviation of Total Reserves:

Independent	117,026
Correlated Factors, Indep AYs	136,698
Correlated AYs	205,845

Comparison of Indications

- Simple method sometimes used -Standard Deviation of Indications
- Implicit assumptions that each method error has the same standard deviation σ , and that the method errors are uncorrelated
- Let L_i be the indicated ultimate loss for method i .

$$S = \frac{\sum_{i=1}^n (L_i - \bar{L})^2}{n - 1}$$

Independent Indications

- $E(S) = \sigma$
- If the selected loss is a straight average of the individual indications, the final estimate would actually have a standard deviation of $\frac{\sigma}{\sqrt{n}}$ with an estimate provided by $\frac{S}{\sqrt{n}}$

Correlated Indications

- $E(S) = \sigma \sqrt{1 - \text{Avg}(\rho_{i,j} | i \neq j)}$
- Example of indication correlation:

	Incurred			
	Paid LDF	LDF	Paid BF	Incurred BF
Paid LDF	1	0.5	0.75	0.4
Incurred LDF	0.5	1	0.5	0.85
Paid BF	0.75	0.5	1	0.65
Incurred BF	0.4	0.85	0.65	1

- $S = 0.63 * \sigma$

Correlated Indications

- Additionally the standard deviation of a straight average of the indications is different:

$$\sigma \sqrt{\frac{\sum \rho_{t,j}}{n^2}}$$

- In our example this is 0.70σ

	True Amount	Estimate (Expected Value)
Individual Method	σ	0.63σ
Average	0.70σ	$0.63 * s / 2 = .37 \sigma$

Problems with Skewed Distributions

- Simple example – Small excess general liability triangle with limited or no losses to date
- You know the potential for significant losses is there, but it so far it hasn't happened.
- A priori loss ratios and development factors are being used to develop a B-F reserve estimate.
- **Don't** base your estimate of reserve variability on historical deviations from the **historical** mean
- Base your estimate of reserve variability on historical deviations from the **expected** mean

Tail Factors

- Large tail factors represent significant uncertainty.
- Even if these factors are based on historical observation of larger triangles, the data is old and the environment has changed.
- Estimating the standard deviation of tail development by comparison with similarly sized factors earlier in the development pattern is likely to understate the variability.
- Consider making adjustments, for example assuming that tail development is perfectly correlated with the development that comes before it.

The Role for Judgment

- Informed judgment plays an important role in most actuarial analyses.
- Most actuaries are comfortable with using judgment when selecting reserve estimates (i.e. means)
- Since estimating reserve variability can be challenging, judgment is important here as well.
- How have you seen loss development estimates change in the past? How comfortable do you feel about the analysis including, any inputs such as reference patterns, etc.
- Put yourself in the position of someone personally guaranteeing the results to remove the process from being purely technical.