Aggregation of results from stochastic reserving methods: Worked example

David Hindley

October 10, 2017

Abstract

This document summarises a worked example of the aggregation of results across classes of business, produced using stochastic reserving methods. It explores the relationship between correlation assumptions, copula approaches and diversification benefits.

1 Introduction

This article describes a worked example of the aggregation of results across classes of business, produced using stochastic reserving methods. It is designed to demonstrate the impact on results of using different assumptions for each class of business and to explore how different correlation assumptions affect, for example, the diversification benefits produced when the results are aggregated across classes. Results are shown using a simple "Variance/Covariance" approach as well as different copula approaches. The web-based software application that accompanies the book entitled "Claims Reserving in General Insurance", by David Hindley, includes a module that allows the user to reproduce the results in this article, and to explore the impact on results of alternative assumptions. This application is available at https://goo.gl/ZrEcM5. All references to "Sections" in this article are to sections in that book.

This article is in draft form at present. Please send any comments to info@claimsreserving.com.

2 Assumptions

For this simplified example, it is assumed that there are three classes of business where a mathematical distribution has been selected for the total future claims across all cohorts.

In practice, if, for example, bootstrapping had been used for each class of business, then the bootstrap simulation results could be combined instead to produce the aggregate distribution, using, for example, a copula re-sorting approach, as outlined in Section 4.9.1. If this re-sorting approach was applied, for example, to the distribution of total future claims across all cohorts, then the resulting sort order could also be used for other components of each simulation that might be of interest, such as the individual cohorts or CDR, thus implying the same dependency between classes as for the total across cohorts.

The use of mathematical distributions, rather than bootstrapping in this example makes the implementation of the aggregation straightforward using the R software. This then enables the main focus of the worked example to be on the conclusions that can be drawn from the results using different correlation and copula assumptions. The same observations would apply equally well if bootstrap simulation results had been aggregated.

The distributional and associated parameter assumptions for each of the three classes are shown in Table 1. The appendix to this article discusses the derivation of the Lognormal parameters, including some additional material to supplement that given in Section 6.9.4 of the reserving book.

It is assumed that both parameter and process error have been allowed for in these distributions, and hence the standard deviation assumption is referred to as the Prediction Error or "PE". Class A is based on the same Taylor and Ashe data that has been used in the worked examples in the reserving book, and has the same reserve as derived by applying Mack's method to this data, including a tail factor, as shown in Reserving book Table 4.20. The CV has been selected as 16%, which was derived by applying the ODP Bootstrap (with constant scale parameter) to the data, as shown in Reserving book Table 4.20 [1].

Table 1
Aggregation example: Summary of individual classes of business

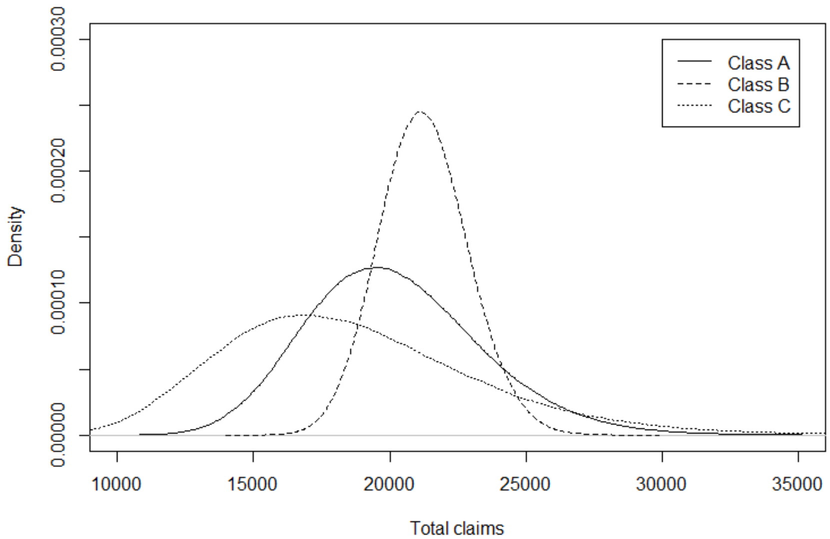
Class	Reserve	Pred'n Error	CV	Fitted Dist'n	Parameters
A	20,219	3,235	16%	Lognormal	$\mu = 9.9017; \sigma = 0.1589$
B	21,250	1,630	8%	Gamma	$\beta = 125; \gamma = 170$
C	18,606	4,725	25%	Lognormal	$\mu = 9.8; \sigma = 0.25$
Total	60,075				

Class B and C have been chosen to have lower and higher CV's respectively, when compared to Class A. The relative variability of the future claims distribution for each class can also be seen in Figure 1, which shows the distribution of the three classes. The lower variability of Class B is clearly evident from this figure.

Another way of comparing the three distributions is to compare the percentile values. Table 2 shows the percentile values and the implied ratio of these values to the mean reserve for each class. The total column in this table

[1] The use of 16% rather than the value produced from the Mack method (of 13%) is just to demonstrate that in some cases, judgement might be used to blend the results from different methods.

Figure 1
Aggregation example:Distribution of three classes



assumes that the future claims for the three classes of business are fully correlated - i.e. that there is no diversification benefit between them. Hence, the total percentile values are the sum of those for the individual classes. The PE is the same as the fully correlated value in the Variance-Covariance approach, as explained in that subsection below.

For Class A, the ratio of the values at different percentile levels to the mean, as shown in Table 2, are very similar[2] to those produced from the ODP Bootstrap process for the Taylor and Ashe data, as shown in Reserving book Table 4.43, suggesting that the Lognormal is a reasonable fit to the bootstrap data.

For this example, to understand the impact on the aggregate distribution of different dependency assumptions between the classes, results have been derived using the following:

- Case 1: Simple linear correlation using the Variance/Covariance approach.
- Case 2: Full independence.
- Case 3: Gaussian copula.
- Case 4: t-copula with 1 degree of freedom[3].

[2] In fact, they are identical at all the stated percentiles below 99.5^{th}, and only marginally different at that level.

[3] Note that the $t-$copula with 1 degree of freedom is equivalent to the Cauchy copula.

Table 2
Aggregation example: Distributions of individual classes

Item	Class A	Class B	Class C	Total[a]
Mean	20,219	21,250	18,606	60,075
Pred' Error	3,235	1,630	4,725	9,590
CV	16%	8%	25%	16%

	Value for Class:			
Percentile	A	B	C	Total[a]
50.00%	19,965	21,208	18,034	59,207
75.00%	22,225	22,325	21,346	65,896
90.00%	24,477	23,364	24,844	72,685
95.00%	25,932	24,000	27,206	77,139
99.50%	30,069	25,682	34,336	90,088
99.90%	32,632	26,645	39,048	98,324

	Ratio to mean for Class:			
Percentile	A	B	C	Total[a]
50.00%	99%	100%	97%	99%
75.00%	110%	105%	115%	110%
90.00%	121%	110%	134%	121%
95.00%	128%	113%	146%	128%
99.50%	149%	121%	185%	150%
99.90%	161%	125%	210%	164%

[a] Undiversified

- Case 5: t-copula with 4 degrees of freedom.

To demonstrate the interaction of choice of copula and level of correlation, as well as Case 2, with no correlation, a High and a Low scenario for the correlation between the classes has been assumed for each of the other cases:

$$\text{High} = \begin{array}{c c} \text{Class} & \begin{array}{ccc} {\scriptstyle 1} & {\scriptstyle 2} & {\scriptstyle 3} \end{array} \\ \begin{array}{c} {\scriptstyle 1} \\ {\scriptstyle 2} \\ {\scriptstyle 3} \end{array} & \begin{bmatrix} 1.0 & 0.8 & 0.6 \\ 0.8 & 1.0 & 0.5 \\ 0.6 & 0.5 & 1.0 \end{bmatrix} \end{array} ; \text{Low} = \begin{array}{c c} \text{Class} & \begin{array}{ccc} {\scriptstyle 1} & {\scriptstyle 2} & {\scriptstyle 3} \end{array} \\ \begin{array}{c} {\scriptstyle 1} \\ {\scriptstyle 2} \\ {\scriptstyle 3} \end{array} & \begin{bmatrix} 1.0 & 0.1 & 0.2 \\ 0.1 & 1.0 & 0.1 \\ 0.2 & 0.1 & 1.0 \end{bmatrix} \end{array} \quad (1)$$

For the purpose of this simplified example, it is assumed that these are linear correlations. In practice, if rank correlations had been determined, then they can be transformed into linear correlations, as referred to in the Gaussian copula procedure given in Section 4.9.1.

3 Results - Case 1: Variance/Covariance approach

Here, the distribution assumptions in Table 1 are not needed, as the Prediction Errors are just combined using the procedure outlined in Section 4.9.2. The $\boldsymbol{S}$ matrix referred to in that section will include the Prediction Errors from Table 1 on its diagonal, that is:

$$\boldsymbol{S} = \begin{pmatrix} 3,235 & 0 & 0 \\ 0 & 1,630 & 0 \\ 0 & 0 & 4,725 \end{pmatrix}$$

Then, for the Low correlation example, this is combined with the correlation matrix above, which is denoted by $\boldsymbol{V}$, to give:

$$\boldsymbol{SVS} = \begin{pmatrix} 10,465,225 & 527,240 & 3,057,195 \\ 527,240 & 2,656,250 & 770,111 \\ 3,057,195 & 770,111 & 22,327,373 \end{pmatrix} \tag{2}$$

The square root of the sum of these values then represents the estimated Prediction Error for the aggregate reserve. This is 6,645, which gives a CV of 11% using the mean reserve of 60,075. The same matrix calculation is done using the High correlation matrix, as well as with the correlations all set to one, and the opposite - fully independent - is calculated by taking the square root of the sum of the individual PE's. A summary of the results is given in Table 3. For

Table 3
Aggregation example: Variance/Covariance approach - Summary of Standard Errors of Aggregate distribution

Correlation	PE	CV
Zero	5,954	9.9%
Low	6,645	11.1%
High	8,362	13.9%
Full	9,590	16.0%

illustration purposes only, if it is assumed that the Lognormal distribution is appropriate for the aggregate distribution of claims across all classes combined, then each of these PEs (along with the mean of 60,075) can be used to estimate the Lognormal parameters and then the reserves at selected percentile values (using the approach given in Section 6.9.4). These are shown in the "Case 1 - VCV" columns in Tables 4 and 5, which compare the results for all the cases and correlation assumptions.

4 Results - Cases 2 to 4: Independence and Copula approach

To produce the results for these cases, the R software package was used, as it has very convenient copula functions, as well as the ability to model a wide range of marginal distributions for each reserving category. Other software packages could of course also be used. The relevant R code is included in the software application referred to in the abstract. Rather than provide the detailed workings here, which would be impractical for such complex calculations, a summary of the algorithm that is used in R is outlined below.

- Define the three marginal distributions for the aggregate claims of each class and their associated parameters, as per Table 1.
- Calculate summary statistics and percentiles for these three distributions, as shown in Table 2.
- If required, draw a plot of the distributions, as shown in Figure 1, just to check that they look reasonable.
- Specify the correlation matrices, as per those shown in the Variance/ Covariance case.
- Specify the types of copulas that will be used (these will be Independent, Gaussian, $t-$copula 1df and $t-$ copula 4df in this example).
- Create the aggregate (or "joint") distribution using the selected marginal distributions along with each combination of copula type and correlation assumption. In this example, this was implemented using the "copula" package in the R software.
- Simulate from each of these joint distributions (100,000 values were simulated for this example).
- If required, draw scatter-plots of the simulated results, showing, for selected pairs of classes, how they are related under each dependency assumption (e.g. to compare tail-dependency).
- Determine statistics for these joint distributions, including percentile values.
- If required, draw plots of the different aggregate cumulative distributions to compare results graphically.

The numerical results are summarised in Tables 4 and 5. "Case 0" in this table is simply the result of adding the values for each class, without any diversification credit - as per the Total column in Table 1. The diversification credit implied by the Gaussian copula and the two $t-$copulas is summarised in Table 6, which shows the reduction in the reserve value, compared to the undiversified results (as a % of the undiversified results of Case 0). Figures 2 and 3 show some example scatter-plots for different copula and correlation assumptions - in each case for the pair of classes A and B.

Table 4
Aggregation example: Comparison of results for different dependency approaches - Part 1

	Case 0 (Undivers'd)	Case 1(VCV) Correlation:				Case 2(Indep')	Case 3(Gaussian) Correlation:		
	-	Zero	Low	High	Full	-	Zero	Low	High
Mean	60,075	60,075	60,075	60,075	60,075	60045	60,083	60,083	60,080
PE	9,590	5,954	6,645	8,362	9,590	5948	5,962	6,642	8,343
CV	16.0%	9.9%	11.1%	13.9%	16.0%	9.9%	9.9%	11.1%	13.9%
50.00%	59,207	59,782	59,711	59,501	59,324	59,604	59,642	59,573	59,403
75.00%	65,896	63,905	64,322	65,329	66,023	63,711	63,781	64,202	65,219
90.00%	72,685	67,857	68,775	71,061	72,697	67,781	67,859	68,749	71,009
95.00%	77,139	70,339	71,586	74,729	77,010	70,442	70,533	71,738	74,787
99.50%	90,088	77,120	79,326	85,015	89,266	78,398	78,483	80,690	85,870
99.90%	98,324	81,144	83,956	91,295	96,855	83,523	83,536	86,235	92,596

Table 5
Aggregation example: Comparison of results for different dependency approaches - Part 2

	Case 4 ($t-$copula 1df) Correlation:			Case 5 ($t-$copula 4df) Correlation:		
	Zero	Low	High	Zero	Low	High
Mean	60,069	60,067	60,061	60,084	60,084	60,081
PE	5,989	6,583	8,157	5,964	6,631	8,322
CV	10.0%	11.0%	13.6%	9.9%	11.0%	13.9%
50.00%	59,642	59,601	59,514	59,644	59,590	59,427
75.00%	62,896	63,283	64,358	63,487	63,909	65,002
90.00%	66,984	67,869	70,239	67,498	68,390	70,712
95.00%	70,442	71,608	74,524	70,398	71,630	74,712
99.50%	82,476	83,981	87,652	80,090	82,136	87,033
99.90%	90,300	91,613	95,485	86,680	88,979	94,579

Several observations can be made regarding the numerical results. Some of these will apply in most situations and are very intuitive and somewhat obvious, but they are stated anyway for completeness. Others may only apply to this specific example.

- The mean values for the reserves are very slightly different to the theoretical mean for the aggregate distribution (of 60,075) in some cases, which is due to the fact that simulation is used to produce the distribution of

Table 6
Aggregation example: Summary of diversification benefits

	Case 3 - Gaussian			Case 4 - $t-$copula 1df			Case 5 - $t-$copula 4df		
Percentile	Zero	Low	High	Zero	Low	High	Zero	Low	High
50.00%	-0.73%	-0.62%	-0.33%	-0.73%	-0.67%	-0.52%	-0.74%	-0.65%	-0.37%
75.00%	3.21%	2.57%	1.03%	4.55%	3.97%	2.33%	3.66%	3.02%	1.36%
90.00%	6.64%	5.41%	2.31%	7.84%	6.63%	3.36%	7.14%	5.91%	2.71%
95.00%	8.56%	7.00%	3.05%	8.68%	7.17%	3.39%	8.74%	7.14%	3.15%
99.50%	12.88%	10.43%	4.68%	8.45%	6.78%	2.70%	11.10%	8.83%	3.39%
99.90%	15.04%	12.30%	5.83%	8.16%	6.83%	2.89%	11.84%	9.50%	3.81%

Figure 2
Aggregation example: Scatter plots of Total reserves for different copula

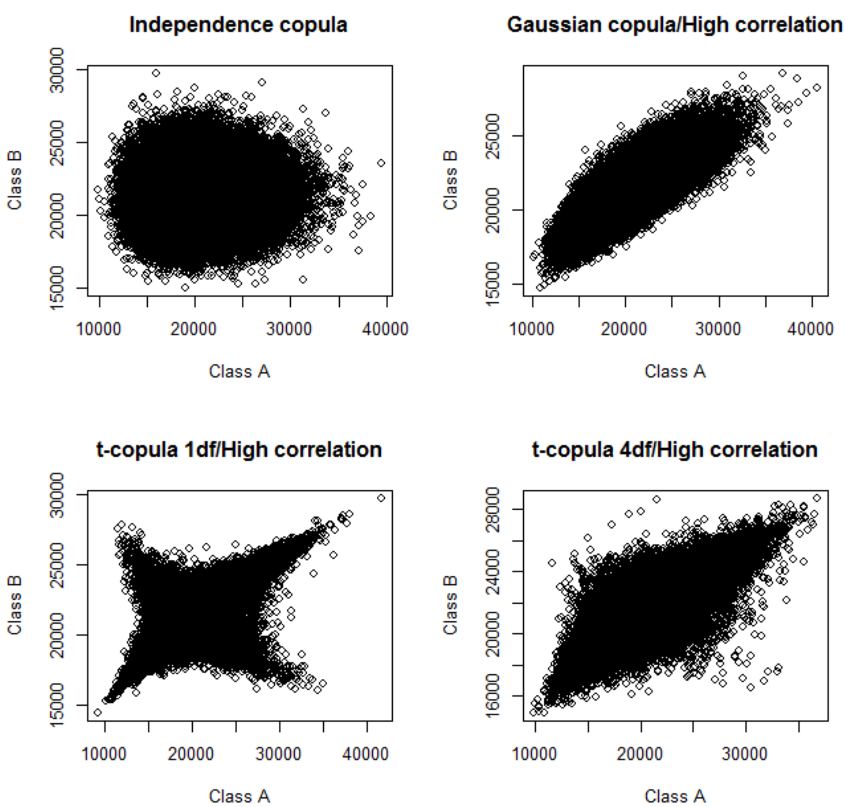
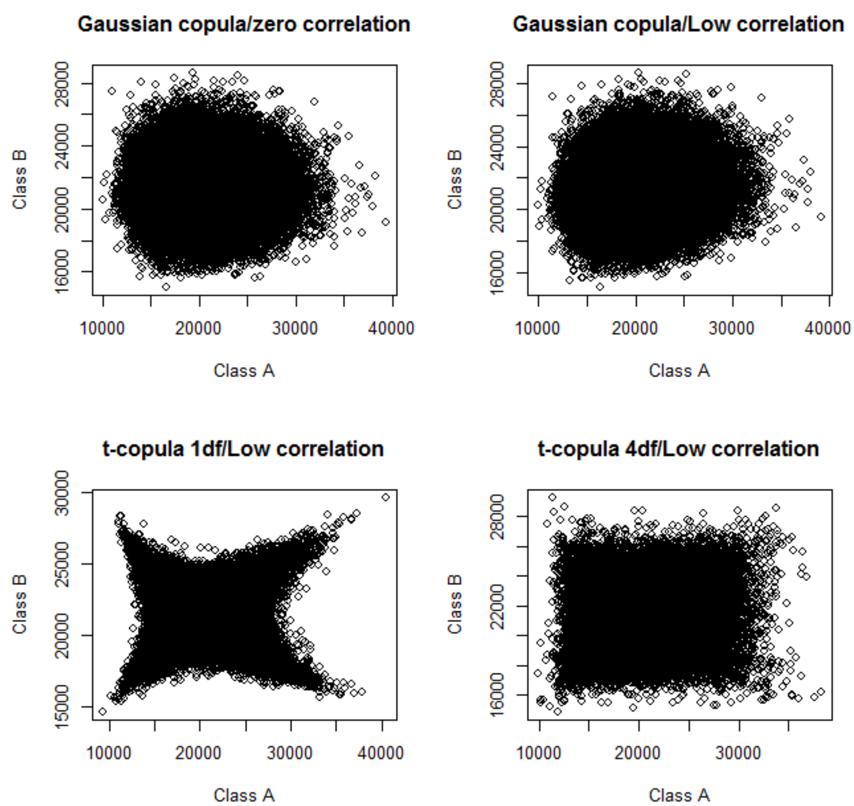


Figure 3
Aggregation example: Scatter plots of Total reserves for different copula - Continued



the aggregate claims (using the relevant copula functionality in the R software). If bootstrap results were being re-sorted instead, then the means would all be the same.

- For all percentiles above 50%, the higher the correlation, the higher the estimated future claims. In other words, the higher the correlation, the lower the diversification benefit.
- For each correlation and dependency assumption, the higher the percentile, the higher the proportionate diversification benefit (except for the t-copula with 1df, where there appears to be a discontinuity at the 95th percentile - perhaps due to the fact that simulation is used).
- The variability(as measured by the CV) increases as the level of correlation increases.
- The VCV approach produces a reasonable approximation to the Gaussian copula, for each correlation assumption, although in this case it seems to slightly overstate the results for high and full correlation, but slightly understate the results at the very high percentiles for the zero and low correlation assumptions, and so should be used with caution.
- The results for the Gaussian copula with zero correlation are very close to the results for the Independence copula. In theory they should be the same, but simulation will cause some differences. In practice, the Gaussian copula with zero correlation could possibly be used to produce results assuming that all classes are independent, leading to an effective lower bound for the reserve values at higher percentiles.
- At the higher percentiles, the t-copulas produce higher values than than the Gaussian copula, reflecting the effect of tail-dependency in the t-copulas. In other words, the tail-dependency of the t-copulas reduces the diversification benefit compared to the Gaussian copula with the same correlation assumptions.
- The increase, compared to the Gaussian copula, caused by the tail dependency of the t copulas is reduced the higher the correlations. In other words, if there is low correlation between the classes, the impact of using the t-copula instead of the Gaussian copula, will be proportionately greater than with higher correlation.
- For the t-copulas, even with zero correlation, the values are higher than the Independence copula (and Gaussian with zero correlation) reflecting the fact that there is still some tail-dependency introduced through the use of the t-copula.
- For the t-copulas, in the tail (above the 95th percentile here) for a given correlation assumption the reserves using 4 degrees of freedom are lower

than using one degree of freedom. This reflects the fact that, for the t-copula, there is less tail dependency as the degrees of freedom increases. In other words, in the tail, Case 4 produces less diversification benefit than Case 5, as the degrees of freedom are lower.

- The 1df t-copula results with Low correlation give only slightly more diversification benefit than the Gaussian copula with High correlation. Hence, if in a particular situation, for whatever reason, it was only possible to produce results using the Gaussian copula, then a very approximate proxy for the impact of tail-dependency could be achieved by increasing the correlation assumptions above the baseline assumptions. This should obviously be used with caution, and where tail-dependency needs to be allowed for, it would normally be preferable to find a way to allow for it explicitly through for example the t-copula or through other copulas.

From the scatter-plots given in Figures 2 and 3 it can be seen that:

- The independence copula in Figure 2 is very similar, as expected, to the Gaussian copula with zero correlation in Figure 3, confirming the equivalence of these two approaches.
- The impact of high correlation, compared with low correlation, is seen clearly by comparing the corresponding copula graphs in the two figures.
- The star shape of the $t-$copula can be a feature of this type of copula when the degrees of freedom are low, depending on the selected marginal distributions. The implied relationship between the variables is somewhat complex and may not be a desirable feature.
- With low correlation, the t-copula with higher degrees of freedom is reasonably similar to the Gaussian copula with low correlation, indicating the lessening impact of tail-dependency as the degrees of freedom increases.

Appendix A Deriving parameters for the Lognormal distribution

Recall that the mean and variance of the Lognormal distribution are defined as:

$$\text{Mean} = \exp(\mu + \sigma^2/2) \tag{3}$$

$$\text{Variance} = \exp(2\mu + \sigma^2)(\exp(\sigma^2) - 1) \tag{4}$$

As explained in Section 6.9.4 of the book, where the mean and variance have been estimated (denoted here by R and $p.e.(R)^2$ respectively), the parameters of the Lognormal can be derived using:

$$\sigma^2 = \ln(1 + (p.e.(R)/R)^2) \tag{5}$$

$$\mu = \ln(R) - \sigma^2/2 \tag{6}$$

These two equations can be used to derive the Lognormal parameters for Class A and C shown in Table 1, by putting R as the relevant Reserve value in the table and p.e.(R) as the Prediction Error.

As noted in Section 6.9.4, in some stochastic reserving contexts, rather than estimate the mean and variance, it might be more practical to derive the mean and an estimate of the reserves at a chosen percentile. This might be the case where there is insufficient data to apply a stochastic reserving method, but where it might be possible, for example, to use expert judgement and/or benchmarking to derive an approximate estimate of the reserves at a specified level of confidence. Section 6.9.4 explains how, for example, the ”solver” functionality of some spreadsheet packages can be used to derive the corresponding Lognormal parameters, and hence then an estimate of the reserves at other percentiles. It is worth noting that these parameters can also be derived analytically, as explained further below.

The cumulative density function, $F_z(z)$ of the Lognormal distribution is defined as

$$F_z(z) = \Phi\Big[\frac{\ln(z) - \mu}{\sigma}\Big] \tag{7}$$

Where Φ is the standard cumulative density function of the $N(0,1)$ distribution.

So, if a value of the reserves at the α percentile has been estimated to be z, then:

$$\alpha = \Phi\Big[\frac{\ln(z) - \mu}{\sigma}\Big] \tag{8}$$

This can be rearranged as:

$$\ln(z) = \mu + \sigma\Phi^{-1}(\alpha) \tag{9}$$

or

$$\mu + \sigma\Phi^{-1}(\alpha) = \ln(z) \tag{10}$$

Then, denoting the estimated mean as m, taking logs of (3) gives:

$$\mu + \sigma^2/2 = \ln(m) \tag{11}$$

Next, subtracting (10) from (11) enables μ to be eliminated, and gives:

$$\sigma^2/2 - \sigma\Phi^{-1}(\alpha) = \ln(m) - \ln(z) \tag{12}$$

Multiplying this by 2 gives a quadratic in σ:

$$\sigma^2 - 2\sigma\Phi^{-1}(\alpha) - 2(\ln(m) - \ln(z)) = 0 \tag{13}$$

This can be solved for sigma, using the usual quadratic formula, so that:

$$\sigma = \frac{2B \pm \sqrt{4B^2 + 8[\ln(m) - \ln(z)]}}{2} \tag{14}$$

Where $B = \Phi^{-1}(\alpha)$.

Then, once the two values of σ have been derived, the corresponding values of μ can be derived using either of the two original equations, so that, for example using (6) gives:

$$\mu = \ln(m) - \sigma^2/2 \tag{15}$$

For the simple example in Section 6.9.4, with the mean reserve (i.e. m) equal to 100 and the value at the 75th percentile (i.e. $\alpha = 0.75$) estimated to be 117 (i.e. $z = 117$), using this approach gives two solutions for μ and σ as follows:

$$\mu = 4.054031; \sigma = 1.049895 \text{ and } \mu = 4.56044; \sigma = 0.299085$$

The second of these is the same as the one quoted in the footnote in Section 6.9.4, which produces a value at the 99.5th percentile of 206.6. The first solution produces a much higher value at this percentile of 861.25. In practice, whatever approach is used, if more than one solution is produced, judgement will be needed to decide which is more likely to represent the distribution of future claims. In this simple example, the first solution appears to produces potentially extreme values at the higher percentiles; hence the second solution may be preferable (and is the one produced by the "solver" approach described in Section 6.9.4).

The quadratic formula can also be tested by applying it to Class A in Table 1. Instead of using the estimated mean and prediction error, the mean and a value at a selected percentile are used. For example, assume that the 75th percentile value as shown in Table 2 for Class A had been estimated. Obviously, since the Lognormal parameters that were used to derive this value are as per those in Table 2, applying the quadratic formula should yield the same Lognormal parameter values. This can be verified by using the 75th percentile value as 22,225 (so that $\alpha = 0.75$ and $z = 22,225$) and the mean value (i.e. m) as 20,219 in the quadratic formula, from which one of the solutions does indeed produce the same μ and σ values.