LDF Curve-Fitting and Stochastic Reserving: A Maximum Likelihood Approach

David R. Clark, FCAS, MAAA

LDF Curve-Fitting and Stochastic Reserving: A Maximum Likelihood Approach

or

How to Increase Reserve Variability with Less Data

David R. Clark
American Re-Insurance

2003 Reserves Call Paper Program

Abstract

An application of Maximum Likelihood Estimation (MLE) theory is demonstrated for modeling the distribution of loss development based on data available in the common triangle format. This model is used to estimate future loss emergence, and the variability around that estimate. The value of using an exposure base to supplement the data in a development triangle is demonstrated as a means of reducing variability. Practical issues concerning estimation error and extrapolation are also discussed.

The author gratefully acknowledges the help and encouragement of the following people:
Dick Currie, Jeff Davis, Leigh Halliwell, Don Mango, Dave Spiegler, and Chuck Thayer.

Introduction

Many papers have been written on the topic of statistical modeling of the loss reserving process. The present paper will focus on one such model, making use of the theory of maximum likelihood estimation (MLE) along with the common Loss Development Factor and Cape Cod techniques. After a review of the underlying theory, the bulk of this paper is devoted to a practical example showing how to make use of the techniques and how to interpret the output.

Before beginning a discussion of a formal model of loss reserving, it is worth re-stating the objectives in creating such a model.

The primary objective is to provide a tool that describes the loss emergence (either reporting or payment) phenomenon in simple mathematical terms as a guide to selecting amounts for carried reserves. Given the complexity of the insurance business, it should never be expected that a model will replace a knowledgeable analyst, but the model can become one key indication to assist them in selecting the reserve.

A secondary objective is to provide a means of estimating the range of possible outcomes around the "expected" reserve. The range of reserves is due to both random "process" variance, and the uncertainty in the estimate of the expected value.

From these objectives, we see that a statistical loss reserving model has two key elements:

- The expected amount of loss to emerge in some time period
- The distribution of actual emergence around the expected value

These two elements of our model will be described in detail in the first two sections of this paper. The full paper is outlined as follows:

Section 1:	Expected Loss Emergence
Section 2:	The Distribution of Actual Loss Emergence and Maximum Likelihood
Section 3:	Key Assumptions of the Model
Section 4:	A Practical Example
Section 5:	Comments and Conclusion

The practical example includes a demonstration of the reduction in variability possible from the use of an exposure base in the Cape Cod reserving method. Extensions of the model for estimating variability of the prospective loss projection or of discounted reserves are discussed more briefly.

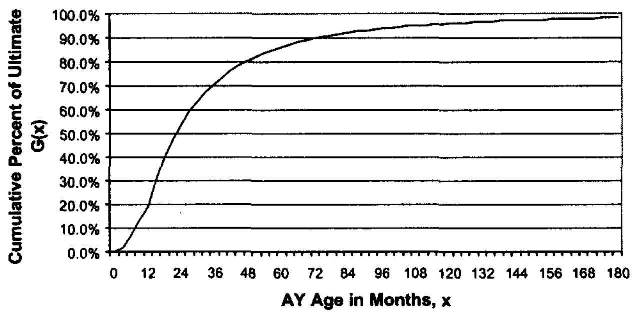
Most of the material presented in this paper makes use of maximum likelihood theory that has already been described more rigorously elsewhere. The mathematics presented here is sufficient for the reader to reproduce the calculations in the examples given, but the focus will be on practical issues rather than on the statistical theory itself.

Section 1: Expected Loss Emergence

Our model will estimate the expected amount of loss to emerge based on a) an estimate of the ultimate loss by year, and b) an estimate of the pattern of loss emergence.

For the expected emergence pattern, we need a pattern that moves from 0 to 100% as time moves from 0 to 8. For our model, we will assume that this pattern is described using the form of a cumulative distribution function[1] (CDF), since a library of such curves is readily available.

$$G(x) \;=\; 1/LDF_x \quad = \text{cumulative \% reported (or paid) as of time x}$$



We will assume that the time index "x" represents the time from the "average" accident date to the evaluation date. The details for approximating different exposure periods (e.g., accident year versus policy year) are given in Appendix B.

For convenience, the model will include two familiar curve forms: Weibull and Loglogistic. Each of these curve forms can be parameterized with a scale θ and a shape ω ("warp"). The Loglogistic curve is familiar to many actuaries under the name "inverse

[1] We are using the <u>form</u> of the distribution function, but do not mean to imply any probabilistic model. The paper by Weissner [9] makes the report lag itself the random variable. By contrast, the loss dollars will be the random variable in our application.

power" (see Sherman[2] [8]), and will be considered the benchmark result. The Weibull will generally provide a smaller "tail" factor than the Loglogistic.

The Loglogistic curve has the form:

$$G(x \mid \omega, \theta) = \frac{x^{\omega}}{x^{\omega} + \theta^{\omega}} \qquad LDF_x = 1 + \theta^{\omega} \cdot x^{-\omega}$$

The Weibull curve has the form:

$$G(x \mid \omega, \theta) = 1 - \exp\left(-(x/\theta)^{\omega}\right)$$

In using these curve forms, we are assuming that the expected loss emergence will move from 0% to 100% in a strictly increasing pattern. The model will still work if some actual points show decreasing losses, but if there is real expected negative development (e.g., lines of business with significant salvage recoveries) then a different model should be used.

There are several advantages to using parameterized curves to describe the expected emergence pattern. First, the estimation problem is simplified because we only need to estimate the two parameters. Second, we can use data that is not strictly from a triangle with evenly spaced evaluation dates – such as the frequent case in which the latest diagonal is only nine months from the second latest diagonal. Third, the final indicated pattern is a smooth curve and does not follow every random movement in the historical age-to-age factors.

The next step in estimating the amount of loss emergence by period is to apply the emergence pattern G(x), to an estimate of the ultimate loss by accident year.

Our model will base the estimate of the ultimate loss by year on one of two methods: either the LDF or the Cape Cod method. The LDF method assumes that the ultimate loss

[2] Sherman actually applies the inverse power curve to the link ratios between ages. Our model will apply this curve to the age-to-ultimate pattern.

amount in each accident year is independent of the losses in other years. The Cape Cod method assumes that there is a known relationship between the amount of ultimate loss expected in each of the years in the historical period, and that this relationship is identified by an exposure base. The exposure base is usually onlevel premium, but can be any other index (such as sales or payroll), which is reasonably assumed to be proportional to expected loss.

The expected loss for a given period will be denoted:

$\mu_{AY;x,y}$ = expected incremental loss dollars in accident year AY between ages x and y

Then the two methods for the expected loss emergence are:

Method #1: "Cape Cod"

$$\mu_{AY;x,y} = \text{Premium}_{AY} \cdot ELR \cdot [G(y \mid \omega,\theta) - G(x \mid \omega,\theta)]$$

Three parameters: ELR, ω, θ

Method #2: "LDF"

$$\mu_{AY;x,y} = ULT_{AY} \cdot [G(y \mid \omega,\theta) - G(x \mid \omega,\theta)]$$

n+2 Parameters: n Accident Years (one ULT for each AY) + ω, θ

While both of these methods are available for use in estimating reserves, Method #1 will generally be preferred. Because we are working with data summarized into annual blocks as a development triangle, there will be relatively few data points included in the model (one data point for each "cell" in the triangle). There is a real problem with overparameterization when the LDF method is used.

For example, if we have a triangle for ten accident years then we have provided the model with 55 data points. The Cape Cod method requires estimation of 3 parameters, but the LDF method requires estimation of 12 parameters.

The Cape Cod method may have somewhat higher process variance estimated, but will usually produce a significantly smaller estimation error. This is the value of the information in the exposure base provided by the user[3]. In short: *the more information that we can give to the model, the smaller the reserve variability due to estimation error.*

The fact that variance can be reduced by incorporating more information into a reserve analysis is, of course, the point of our ironic subtitle: How to Increase Reserve Variability with Less Data. The point is obvious, but also easy to overlook. The reduction in variability is important even to those who do not explicitly calculate reserve ranges because it still guides us towards better estimation methods: lower variance implies a better reserve estimate.

[3] Halliwell [2] provides additional arguments for the use of an exposure index. See especially pages 441 - 443.

Section 2: The Distribution of Actual Loss Emergence and Maximum Likelihood

Having defined the model for the expected loss emergence, we need to estimate the "best" parameters for that model and, as a secondary goal, estimate the variance around the expected value. Both of these steps will be accomplished making use of maximum likelihood theory.

The variance will be estimated in two pieces: process variance (the "random" amount) and parameter variance (the uncertainty in our estimator).

2.1 Process Variance

The curve $G(x \mid \omega, \theta)$ represents the expected loss emergence pattern. The actual loss emergence will have a distribution around this expectation.

We assume that the loss in any period has a constant ratio of variance/mean[4]:

$$\frac{Variance}{Mean} = \sigma^2 \approx \frac{1}{n-p} \cdot \sum_{AY,t}^{n} \frac{(c_{AY,t} - \mu_{AY,t})^2}{\mu_{AY,t}}$$

where p = # of parameters

$c_{AY,t}$ = actual incremental loss emergence

$\mu_{AY,t}$ = expected incremental loss emergence

(this is recognized as being equivalent to a chi-square error term)

For estimating the parameters of our model, we will further assume that the actual incremental loss emergence "c" follows an over-dispersed Poisson distribution. That is, the loss dollars will be a Poisson random variable times a scaling factor equal to σ^2.

[4] This assumption will be tested by analysis of residuals in our example.

Standard Poisson: $\Pr(x) = \dfrac{\lambda^x \cdot e^{-\lambda}}{x!}$ $\quad E[x] = Var(x) = \lambda$

Actual Loss: $c = x \cdot \sigma^2$ $\quad \Pr(c) = \dfrac{\lambda^{c/\sigma^2} \cdot e^{-\lambda}}{(c/\sigma^2)!}$ $\quad E[c] = \lambda \cdot \sigma^2 = \mu$

$$Var(c) = \lambda \cdot \sigma^4 = \mu \cdot \sigma^2$$

The "over-dispersed Poisson" sounds strange when it is first encountered, but it quickly proves to have some key advantages. First, inclusion of the scaling factor allows us to match the first and second moments of any distribution, which gives the model a high degree of flexibility. Second, maximum likelihood estimation exactly produces the LDF and Cape Cod estimates of ultimate, so the results can be presented in a format familiar to reserving actuaries.

The fact that the distribution of ultimate reserves is approximated by a discretized curve should not be cause for concern. The scale factor σ^2 is generally small compared to the mean, so little precision is lost. Also, the use of a discrete distribution allows for a mass point at zero, representing the cases in which no change in loss is seen in a given development increment.

Finally, we should remember that this maximum likelihood method is intended to produce the mean and variance of the distribution of reserves. Having estimated those two numbers, we are still free to switch to a different distribution form when the results are used in other applications.

2.2 The Likelihood Function – Finding the "Best" Parameters

The likelihood function is:

$$\text{Likelihood} = \prod_i \Pr(c_i) = \prod_i \frac{\lambda_i^{c_i/\sigma^2} \cdot e^{-\lambda_i}}{(c_i/\sigma^2)!} = \prod_i \frac{(\mu_i/\sigma^2)^{c_i/\sigma^2} \cdot e^{-\mu_i/\sigma^2}}{(c_i/\sigma^2)!}$$

This can be maximized using the logarithm of the likelihood function:

$$\text{LogLikelihood} = \sum_i \left(c_i/\sigma^2\right)\cdot\ln\left(\mu_i/\sigma^2\right) - \mu_i/\sigma^2 - \ln\left(\left(c_i/\sigma^2\right)!\right)$$

Which is equivalent to maximizing:

$$\boxed{\ell = \sum_i c_i \cdot \ln(\mu_i) - \mu_i} \qquad \text{if } \sigma^2 \text{ is assumed to be known}$$

Maximum likelihood estimators of the parameters are found by setting the first derivatives of the loglikelihood function ℓ equal to zero:

$$\frac{\partial \ell}{\partial ELR} = \frac{\partial \ell}{\partial \theta} = \frac{\partial \ell}{\partial \omega} = 0$$

For “Model #1: Cape Cod”, the loglikelihood function becomes:

$$\ell = \sum_{i,t}\left(c_{i,t}\cdot\ln\left(ELR\cdot P_i\cdot[G(x_t)-G(x_{t-1})]\right) - ELR\cdot P_i\cdot[G(x_t)-G(x_{t-1})]\right)$$

where $c_{i,t}$ = actual loss in accident year i, development period t

P_i = Premium for accident year i

x_{t-1} = beginning age for development period t

x_t = ending age for development period t

$$\frac{\partial \ell}{\partial ELR} = \sum_{i,t}\left(\frac{c_{i,t}}{ELR} - P_i\cdot[G(x_t)-G(x_{t-1})]\right)$$

For $\dfrac{\partial \ell}{\partial ELR} = 0, \qquad ELR = \dfrac{\sum_{i,t} c_{i,t}}{\sum_{i,t} P_i \cdot [G(x_t) - G(x_{t-1})]}$

The MLE estimate for ELR is therefore equivalent to the "Cape Cod" Ultimate. It can be set based on θ and ω, and so reduce the problem to be solved to two parameters instead of three.

For "Model #2: LDF", the loglikelihood function becomes:

$$\ell = \sum_{i,t} \left(c_{i,t} \cdot \ln\left(ULT_i \cdot [G(x_t) - G(x_{t-1})]\right) - ULT_i \cdot [G(x_t) - G(x_{t-1})] \right)$$

$$\frac{\partial \ell}{\partial ULT_i} = \sum_t \left(\frac{c_{i,t}}{ULT_i} - [G(x_t) - G(x_{t-1})] \right)$$

For $\dfrac{\partial \ell}{\partial ULT_i} = 0, \qquad ULT_i = \dfrac{\sum_t c_{i,t}}{\sum_t [G(x_t) - G(x_{t-1})]}$

The MLE estimate for each ULT_i is therefore equivalent to the "LDF Ultimate"[5]. It can also be set based on θ and ω, and to again reduce the problem to be solved to two parameters instead of $n+2$.

A final comment worth noting is that the maximum loglikelihood function never takes the logarithm of the actual incremental development $c_{i,t}$. The model will work even if some of these amounts are zero or negative.

[5] See Mack [5], Appendix A, for a further discussion of this relationship.

2.3 Parameter Variance[6]

The second step is to find the variance in the estimate of the parameters. This is done based on the Rao-Cramer approximation, using the second derivative information matrix I, and is commonly called the "Delta Method" (c.f. Klugman, et al [3], page 67).

The second derivative information matrix for the "Cape Cod Method" is 3x3 and assumes the same ELR for all accident years:

$$I = \begin{bmatrix} \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial ELR^2} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial ELR \partial \omega} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial ELR \partial ?} \\ \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial \omega \partial ELR} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial \omega^2} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial \omega \partial ?} \\ \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial ? \partial ELR} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial ? \partial \omega} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial ?^2} \end{bmatrix}$$

The covariance matrix is calculated using the inverse of the Information matrix:

$$\Sigma = \begin{bmatrix} Var(ELR) & Cov(ELR,\omega) & Cov(ELR,\theta) \\ Cov(\omega, ELR) & Var(\omega) & Cov(\omega,\theta) \\ Cov(\theta, ELR) & Cov(\theta,\omega) & Var(\theta) \end{bmatrix} \geq -\sigma^2 \cdot I^{-1}$$

The scale factor σ^2 is again estimated as above:

$$\sigma^2 \approx \frac{1}{n-p} \cdot \sum_{AY,t}^{n} \frac{(c_{AY,t} - \hat{\mu}_{AY,t})^2}{\hat{\mu}_{AY,t}}$$

The second derivative matrix for "LDF Method" is (n+2)x(n+2) and assumes that there is a different ULT for each accident year. The information matrix, I, is given as:

---

[6] To be precise, we are calculating the variance in the estimator of the parameter; the parameter itself does not have any variance. Nonetheless, we will retain the term "parameter variance" as shorthand.

$$\left[\begin{array}{cccc|cc}
\sum_t \frac{\partial^2 \ell_{1,t}}{\partial ULT_1^2} & 0 & \cdots & 0 & \sum_t \frac{\partial^2 \ell_{1,t}}{\partial ULT_1 \partial\omega} & \sum_t \frac{\partial^2 \ell_{1,t}}{\partial ULT_1 \partial ?} \\
0 & \sum_t \frac{\partial^2 \ell_{2,t}}{\partial ULT_2^2} & \cdots & 0 & \sum_t \frac{\partial^2 \ell_{2,t}}{\partial ULT_2 \partial\omega} & \sum_t \frac{\partial^2 \ell_{2,t}}{\partial ULT_2 \partial ?} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & \sum_t \frac{\partial^2 \ell_{n,t}}{\partial ULT_n^2} & \sum_t \frac{\partial^2 \ell_{n,t}}{\partial ULT_n \partial\omega} & \sum_t \frac{\partial^2 \ell_{n,t}}{\partial ULT_n \partial ?} \\
\hline
\sum_t \frac{\partial^2 \ell_{1,t}}{\partial\omega \partial ULT_1} & \sum_t \frac{\partial^2 \ell_{2,t}}{\partial\omega \partial ULT_2} & \cdots & \sum_t \frac{\partial^2 \ell_{n,t}}{\partial\omega \partial ULT_n} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial\omega^2} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial\omega \partial ?} \\
\sum_t \frac{\partial^2 \ell_{1,t}}{\partial ? \partial ULT_1} & \sum_t \frac{\partial^2 \ell_{2,t}}{\partial ? \partial ULT_2} & \cdots & \sum_t \frac{\partial^2 \ell_{n,t}}{\partial ? \partial ULT_n} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial ? \partial\omega} & \sum_{y,t} \frac{\partial^2 \ell_{y,t}}{\partial ?^2}
\end{array}\right]$$

The covariance matrix Σ is again calculated using the inverse of the Information matrix, but for the LDF Method this matrix is larger.

2.4 The Variance of the Reserves

The final step is to estimate the variance in the reserves. The variance is broken into two pieces: the process variances and the estimation error (loosely "parameter variance"). For an estimate of loss reserves R for a given period $\mu_{AY;x,y}$, or group of periods $\sum \mu_{AY;x,y}$, the process variance is given by:

Process Variance of R: $\quad \sigma^2 \cdot \sum \mu_{AY;x,y}$

The estimation error makes use of the covariance matrix Σ calculated above:

Parameter Variance of R: $\quad Var(E[R]) = (\partial R)' \cdot \Sigma \cdot (\partial R)$

where

$$\partial R = \left\langle \frac{\partial R}{\partial ELR}, \frac{\partial R}{\partial \theta}, \frac{\partial R}{\partial \omega} \right\rangle \qquad \text{or} \qquad \partial R = \left\langle \left\{ \frac{\partial R}{\partial ULT_i} \right\}_{i=1}^{n}, \frac{\partial R}{\partial \theta}, \frac{\partial R}{\partial \omega} \right\rangle$$

The future reserve R, under the Cape Cod method is given by:

$$\text{Reserve:} \qquad R = \sum \text{Premium}_i \cdot ELR \cdot \left(G(y_i) - G(x_i)\right)$$

The derivatives needed are then easily calculated:

$$\frac{\partial R}{\partial ELR} = \sum \text{Premium}_i \cdot \left(G(y_i) - G(x_i)\right)$$

$$\frac{\partial R}{\partial \theta} = \sum \text{Premium}_i \cdot ELR \cdot \left(\frac{\partial G(y_i)}{\partial \theta} - \frac{\partial G(x_i)}{\partial \theta}\right)$$

$$\frac{\partial R}{\partial \omega} = \sum \text{Premium}_i \cdot ELR \cdot \left(\frac{\partial G(y_i)}{\partial \omega} - \frac{\partial G(x_i)}{\partial \omega}\right)$$

For the LDF Method, let $\text{Premium}_i = 1$ and $ELR = ULT_i$.

All of the mathematics needed for the estimate of the process and parameter variance is provided in Appendix A. For the two curve forms used, all of the derivatives are calculated analytically, without the need for numerical approximations.

Section 3: Key Assumptions of this Model

- Incremental losses are independent and identically distributed (iid)

  The assumption that all observed points are independent and identically distributed is the famous "iid" of classical statistics. In introductory textbooks this is often illustrated by the problem of estimating the proportion of red and black balls in an urn based on having "randomly" selected a sample from the urn. The "independence" assumption is that the balls are shaken up after each draw, so that we do not always pull out the same ball each time. The "identically distributed" assumption is that we are always taking the sample from the same urn.

  The "independence" assumption in the reserving context is that one period does not affect the surrounding periods. This is a tenuous assumption but will be tested using residual analysis. There may in fact be positive correlation if all periods are equally impacted by a change in loss inflation. There may also be negative correlation if a large settlement in one period replaces a stream of payments in later periods.

  The "identically distributed" assumption is also difficult to justify on first principles. We are assuming that the emergence pattern is the same for all accident years; which is clearly a gross simplification from even a rudimentary understanding of insurance phenomenon. Different risks and mix of business would have been written in each historical period, and subject to different claims handling and settlement strategies. Nonetheless, a parsimonious model requires this simplification.

- The Variance/Mean Scale Parameter σ^2 is fixed and known

  In rigorous maximum likelihood theory, the variance/mean scale parameter σ^2 should be estimated simultaneously with the other model parameters, and the variance around its estimate included in our covariance matrix.

Unfortunately, including the scale parameter in the curve-fitting procedure leads to mathematics that quickly becomes intractable. Treating the scale parameter as fixed and known is an approximation made for convenience in the calculation, and the results are sometimes called "quasi-likelihood estimators". McCullough & Nelder [7] give support for the approximation that we are using.

In effect, we are ignoring the variance on the variance.

In classical statistics, we usually relax this assumption (e.g., in hypothesis testing) by using the Student-T distribution instead of the Normal distribution. Rodney Kreps' paper [4] provides additional discussion on how reserve ranges could increase when this additional source of variability is considered.

- Variance estimates are based on an approximation to the Rao-Cramer lower bound.

The estimate of variance based on the information matrix is only exact when we are using linear functions. In the case of non-linear functions, including our model, the variance estimate is a Rao-Cramer lower bound.

Technically, the Rao-Cramer lower bound is based on the true expected values of the second derivative matrix. Since we are using approximations that plug in the estimated values of the parameters, the result is sometimes called the "observed" information matrix rather than the "expected" information matrix. Again, this is a limitation common to many statistical models and is due to the fact that we do not know the true parameters.

All of the key assumptions listed above need to be kept in mind by the user of a stochastic reserving model. In general, they imply that there is potential for more variability in future loss emergence than the model itself produces.

Such limitations should not lead the user, or any of the recipients of the output, to disregard the results. We simply want to be clear about what sources of variability we are able to measure and what sources cannot be measured. That is a distinction that should not be lost.

Section 4: A Practical Example

4.1 The LDF Method

For the first part of this example, we will use the "LDF Method" (referred to above as "Method 2"). The improvements in the model by moving to the Cape Cod method will be apparent as the numbers are calculated.

The triangle used in this example is taken from the 1993 Thomas Mack paper [6]. The accident years have been added to make the display appear more familiar.

	12	24	36	48	60	72	84	96	108	120
1991	357,848	1,124,788	1,735,330	2,182,708	2,745,596	3,319,994	3,466,336	3,606,286	3,833,515	3,901,463
1992	352,118	1,236,139	2,170,033	3,353,322	3,799,067	4,120,063	4,647,867	4,914,039	5,339,085	
1993	290,507	1,292,306	2,218,525	3,235,179	3,985,995	4,132,918	4,628,910	4,909,315		
1994	310,608	1,418,858	2,195,047	3,757,447	4,029,929	4,381,982	4,588,268			
1995	443,160	1,136,350	2,128,333	2,897,821	3,402,672	3,873,311				
1996	396,132	1,333,217	2,180,715	2,985,752	3,691,712					
1997	440,832	1,288,463	2,419,861	3,483,130						
1998	359,480	1,421,128	2,864,498							
1999	376,686	1,363,294								
2000	344,014									

The incremental triangle, calculated by taking differences between cells in each accident year, is given by:

	12	24	36	48	60	72	84	96	108	120
1991	357,848	766,940	610,542	447,378	562,888	574,398	146,342	139,950	227,229	67,948
1992	352,118	884,021	933,894	1,183,289	445,745	320,996	527,804	266,172	425,046	
1993	290,507	1,001,799	926,219	1,016,654	750,816	146,923	495,992	280,405		
1994	310,608	1,108,250	776,189	1,562,400	272,482	352,053	206,286			
1995	443,160	693,190	991,983	769,488	504,851	470,639				
1996	396,132	937,085	847,498	805,037	705,960					
1997	440,832	847,631	1,131,398	1,063,269						
1998	359,480	1,061,648	1,443,370							
1999	376,686	986,608								
2000	344,014									

This incremental triangle is actually better arranged as a table of values, rather than in the familiar triangular format (see Table 1.1). In the tabular format, the column labeled "Increment" is the value that we will be approximating with the expression

$$\mu_{AY;x,y} \quad = \quad ULT_{AY} \cdot [G(y \mid \omega,\theta) - G(x \mid \omega,\theta)].$$

The x and y values are the "From" and "To" dates.

Before calculating the fitted values, it is worth showing the flexibility in this format. First, if we have only the latest three evaluations of the triangle, we can still use this method directly.

The original triangle becomes:

	12	24	36	48	60	72	84	96	108	120
1991								3,606,286	3,833,515	3,901,463
1992							4,647,867	4,914,039	5,339,085	
1993						4,132,918	4,628,910	4,909,315		
1994					4,029,929	4,381,982	4,588,268			
1995				2,897,821	3,402,672	3,873,311				
1996			2,180,715	2,985,752	3,691,712					
1997		1,288,463	2,419,861	3,483,130						
1998	359,480	1,421,128	2,864,498							
1999	376,686	1,363,294								
2000	344,014									

and the incremental triangle is:

	12	24	36	48	60	72	84	96	108	120
1991								3,606,286	227,229	67,948
1992							4,647,867	266,172	425,046	
1993						4,132,918	495,992	280,405		
1994					4,029,929	352,053	206,286			
1995				2,897,821	504,851	470,639				
1996			2,180,715	805,037	705,960					
1997		1,288,463	1,131,398	1,063,269						
1998	359,480	1,061,648	1,443,370							
1999	376,686	986,608								
2000	344,014									

The tabular format then collapses from 55 rows down to 27 rows, as shown in Table 1.2.

Another common difficulty in working with development triangles is the use of irregular evaluation periods. For example, we may have accident years evaluated at each year-end

- producing ages 12, 24, 36, etc – but the most recent diagonal is only available as of the end of the third quarter (ages 9, 21, 33, etc). This is put into the tabular format by simply changing the evaluation age fields ("Diag Age") as shown in Table 1.3.

Returning to the original triangle, we calculate the fitted values for a set of parameters ULT_{AY}, ω, θ and the MLE term to be maximized.

Fitted Value: $\mu_{AY;x,y} = ULT_{AY} \cdot [G(y \mid \omega,\theta) - G(x \mid \omega,\theta)]$

MLE Term: $c_{AY;x,y} \cdot \ln(\mu_{AY;x,y}) - \mu_{AY;x,y}$

In Table 1.4, these numbers are shown as additional columns. These values also have the desired unbiased property that the sum of the actual incremental dollars $c_{AY;x,y}$ equals the sum of the fitted values $\hat{\mu}_{AY;x,y}$.

The fitted parameters for the Loglogistic growth curve are:

ω	1.434294
θ	48.6249

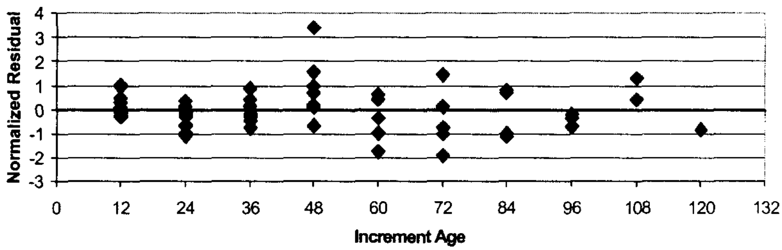
The fitted parameters are found by iteration, which can easily be accomplished in the statistics capabilities of most software packages. Once the data has been arranged in the tabular format, the curve-fitting can even be done in a spreadsheet.

The scale parameter σ^2 is also easily calculated. We recall that the form of this calculation is the same as a Chi-Square statistic, with 43 degrees of freedom (55 data points minus 12 parameters). The resulting σ^2 is 65,029. This scale factor may be thought of as the size of the discrete intervals for the over-dispersed Poisson, but is better thought of simply as the process variance-to-mean ratio. As such, we can calculate the process variance of the total reserve, or any sub-segment of the reserve, by just multiplying by 65,029.

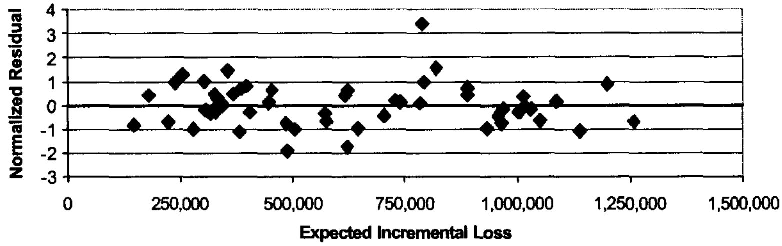
The scale factor σ^2 is also useful for a review of the model residuals (error terms).

$$\text{Normalized Residual:} \qquad r_{AY;x,y} = \frac{\left(c_{AY;x,y} - \hat{\mu}_{AY;x,y}\right)}{\sqrt{\sigma^2 \cdot \hat{\mu}_{AY;x,y}}}$$

The residuals can be plotted in various ways in order to test the assumptions in the model. The graph below shows the residuals plotted against the increment of loss emergence. We would hope that the residuals would be randomly scattered around the zero line for all of the ages, and that the amount of variability would be roughly constant. The graph below tells us that the curve form is perhaps not perfect for the early 12 and 24 points, but the pattern is not enough to reject the model outright.



A second residual plot of the residuals against the expected loss in each increment (the fitted values) is shown below. This graph is useful as a check on the assumption that the variance/mean ratio is constant. If the variance/mean ratio were not constant, then we would expect to see the residuals much closer to the zero line at one end of the graph.



The residuals can also be plotted against the accident year, the calendar year of emergence (to test diagonal effects), or any other variable of interest. The desired outcome is always that the residuals appear to be randomly scattered around the zero line. Any noticeable pattern or autocorrelation is an indication that the some of the model assumptions are incorrect.

Having solved for the parameters ω and θ, and the derived ultimates by year, we can estimate the needed reserves.

Accident Year	Reported Losses	Age at 12/31/2000	Average Age (x)	Growth Function	Fitted LDF	Ultimate Losses	Estimated Reserves
1991	3,901,463	120	114	77.24%	1.2946	5,050,867	1,149,404
1992	5,339,085	108	102	74.32%	1.3456	7,184,079	1,844,994
1993	4,909,315	96	90	70.75%	1.4135	6,939,399	2,030,084
1994	4,588,268	84	78	66.32%	1.5077	6,917,862	2,329,594
1995	3,873,311	72	66	60.78%	1.6452	6,372,348	2,499,037
1996	3,691,712	60	54	53.75%	1.8604	6,867,980	3,176,268
1997	3,483,130	48	42	44.77%	2.2338	7,780,515	4,297,385
1998	2,864,498	36	30	33.34%	2.9991	8,590,793	5,726,295
1999	1,363,294	24	18	19.38%	5.1593	7,033,659	5,670,365
2000	344,014	12	6	4.74%	21.1073	7,261,205	6,917,191
Total	34,358,090					69,998,708	35,640,618

From this initial calculation, we can quickly see the impact of the extrapolated "tail" factor. Our loss development data only includes ten years of development (out to age 120 months), but the growth curve extrapolates the losses to full ultimate. From this data, the Loglogistic curve estimates that only 77.24% of ultimate loss has emerged as of ten years.

Extrapolation should always be used cautiously. For practical purposes, we may want to rely on the extrapolation only out to some finite point – an additional ten years say.

Accident Year	Reported Losses	Age at 12/31/2000	Average Age (x)	Growth Function	Fitted LDF	Truncated LDF	Losses at 240 mo	Estimated Reserves
		240	234	90.50%	1.1050	1.0000		
1991	3,901,463	120	114	77.24%	1.2946	1.1716	4,570,810	669,347
1992	5,339,085	108	102	74.32%	1.3456	1.2177	6,501,273	1,162,188
1993	4,909,315	96	90	70.75%	1.4135	1.2792	6,279,848	1,370,533
1994	4,588,268	84	78	66.32%	1.5077	1.3644	6,260,358	1,672,090
1995	3,873,311	72	66	60.78%	1.6452	1.4888	5,766,692	1,893,381
1996	3,691,712	60	54	53.75%	1.8604	1.6836	6,215,217	2,523,505
1997	3,483,130	48	42	44.77%	2.2338	2.0215	7,041,021	3,557,891
1998	2,864,498	36	30	33.34%	2.9991	2.7140	7,774,286	4,909,788
1999	1,363,294	24	18	19.38%	5.1593	4.6689	6,365,149	5,001,855
2000	344,014	12	6	4.74%	21.1073	19.1012	6,571,068	6,227,054
Total	34,358,090						63,345,723	28,987,633

As noted above, the process variance for the estimated reserve of 28,987,633 is found by multiplying by the variance-to-mean ratio of 65,029. The process standard deviation around our reserve is therefore 1,372,966 for a coefficient of variation (CV = SD/mean) of about 4.7%.

As an alternative to truncating the tail factor at a selected point, such as age 240, we could make use of a growth curve that typically has a lighter "tail". The mathematics for the Weibull curve is provided for this purpose. An example including a fit of the Weibull curve is shown below.

Accident Year	Reported Losses	Age at 12/31/2000	Average Age (x)	Growth Function	Weibull LDF	Ultimate Losses	Estimated Reserves
1991	3,901,463	120	114	95.01%	1.0525	4,106,189	204,726
1992	5,339,085	108	102	92.54%	1.0806	5,769,409	430,324
1993	4,909,315	96	90	89.00%	1.1237	5,516,376	607,061
1994	4,588,268	84	78	84.01%	1.1904	5,461,745	873,477
1995	3,873,311	72	66	77.14%	1.2963	5,020,847	1,147,536
1996	3,691,712	60	54	67.95%	1.4717	5,433,242	1,741,530
1997	3,483,130	48	42	56.01%	1.7853	6,218,284	2,735,154
1998	2,864,498	36	30	41.19%	2.4277	6,954,204	4,089,706
1999	1,363,294	24	18	23.94%	4.1764	5,693,693	4,330,399
2000	344,014	12	6	6.37%	15.6937	5,398,863	5,054,849
Total	34,358,090					55,572,851	21,214,761

The fitted Weibull parameters θ and ω are 48.88453 and 1.296906, respectively. The lower "tail" factor of 1.0525 (instead of 1.2946 for the Loglogistic) may be more in line with the actuary's expectation for casualty business. The difference between the two curve forms also highlights the danger in relying on a purely mechanical extrapolation formula. The selection of a truncation point is an effective way of reducing the reliance on the extrapolation when the thicker-tailed Loglogistic is used.

The next step is our estimate of the parameter variance.

The parameter variance calculation is more involved than what was needed for process variance. As discussed in Section 2.3, we need to first evaluate the Information Matrix, which contains the second derivatives with respect to all of the model parameters, and so is a 12x12 matrix. The mathematics for all of these calculations is given in Appendix A, and is not difficult to program in most software. For purposes of this example, we will simply show the resulting variances:

Accident Year	Reported Losses	Estimated Reserves	Process Std Dev	CV	Parameter Std Dev	CV	Total Std Dev	CV
1991	3,901,463	669,347	208,631	31.2%	158,088	23.6%	261,761	39.1%
1992	5,339,085	1,162,188	274,911	23.7%	257,205	22.1%	376,471	32.4%
1993	4,909,315	1,370,533	298,537	21.8%	298,628	21.8%	422,260	30.8%
1994	4,588,268	1,672,090	329,749	19.7%	356,827	21.3%	485,860	29.1%
1995	3,873,311	1,893,381	350,891	18.5%	401,416	21.2%	533,160	28.2%
1996	3,691,712	2,523,505	405,094	16.1%	518,226	20.5%	657,768	26.1%
1997	3,483,130	3,557,891	481,005	13.5%	704,523	19.8%	853,064	24.0%
1998	2,864,498	4,909,788	565,047	11.5%	968,806	19.7%	1,121,545	22.8%
1999	1,363,294	5,001,855	570,321	11.4%	1,227,880	24.5%	1,353,867	27.1%
2000	344,014	6,227,054	636,348	10.2%	2,838,890	45.6%	2,909,336	46.7%
Total	34,358,090	28,987,633	1,372,966	4.7%	4,688,826	16.2%	4,885,707	16.9%

From this table, one conclusion should be readily apparent: the parameter variance component is much more significant than the process variance. The chief reason for this is that we have overparameterization of our model; that is, the available 55 data points are really not sufficient to estimate the 12 parameters of the model. The 1994 Zehnwirth paper ([10], p. 512f) gives a helpful discussion of the dangers of overparameterization.

The main problem is that we are estimating the ultimate loss for each accident year independently from the ultimate losses in the other accident years. In effect, we are saying that knowing the ultimate loss for accident year 1999 provides no information about the ultimate loss for accident year 2000. As such, our model is fitting to what may just be "noise" in the differences from one year to the next.

This conclusion is unsettling, because it indicates a high level of uncertainty not just in our maximum likelihood model, but in the chain-ladder LDF method in general.

4.2 The Cape Cod Method

A natural alternative to the LDF Method is the Cape Cod method. In order to move on to this method, we need to supplement the loss development triangle with an exposure base that is believed to be proportional to ultimate expected losses by accident year. A natural candidate for the exposure base is onlevel premium – premium that has been adjusted to a common level of rate per exposure.

Unadjusted historical premium could be used for this exposure base, but the impact of the market cycle on premium is likely to distort the results. We prefer onlevel premium so that the assumption of a constant expected loss ratio (ELR) across all accident years is reasonable.

A further refinement would include an adjustment for loss trend net of exposure trend, so that all years are at the same cost level as well as rate level.

There may be other candidates for the exposure index: sometimes the original loss projections by year are available; the use of estimated claim counts has also been suggested. In practice, even a judgmentally selected index may be used.

For the example in the Mack paper, no exposure base was supplied. For this exercise, we will use a simplifying assumption that premium was $10,000,000 in 1991 and increased by $400,000 each subsequent year.

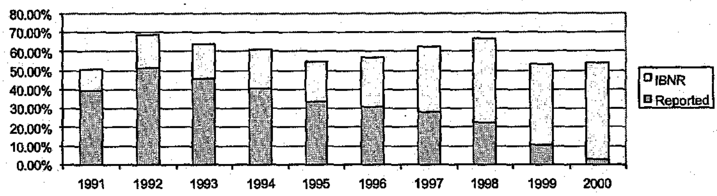
The tabular format of our loss data is shown in Table 2.1. This is very similar to the format used for the LDF Method but instead of the "AY Total" column (latest diagonal), we display the onlevel premium for each accident year. The expected ultimate loss by year is calculated as the ELR multiplied by the onlevel premium.

Accident Year	Onlevel Premium	Age at 12/31/2000	Average Age (x)	Growth Function	Premium x Growth Func	Reported Losses	Ultimate Loss Ratio
1991	10,000,000	120	114	77.76%	7,775,733	3,901,463	50.17%
1992	10,400,000	108	102	74.85%	7,784,279	5,339,085	68.59%
1993	10,800,000	96	90	71.29%	7,699,022	4,909,315	63.77%
1994	11,200,000	84	78	66.87%	7,489,209	4,588,268	61.27%
1995	11,600,000	72	66	61.31%	7,112,024	3,873,311	54.46%
1996	12,000,000	60	54	54.24%	6,508,439	3,691,712	56.72%
1997	12,400,000	48	42	45.17%	5,600,712	3,483,130	62.19%
1998	12,800,000	36	30	33.60%	4,301,252	2,864,498	66.60%
1999	13,200,000	24	18	19.46%	2,568,496	1,363,294	53.08%
2000	13,600,000	12	6	4.69%	638,334	344,014	53.89%
Total	118,000,000				57,477,500	34,358,090	59.78%

The Loglogistic parameters are again solved for iteratively in order to maximize the value of the log-likelihood function in Table 2.1. The resulting parameters are similar to those produced by the LDF method.

ω	1.447634
θ	48.0205

One check that should be made on the data before we proceed with the reserve estimate is a quick test on the assumption that the ELR is constant over all accident years. This is best done with a graph of the estimated ultimate loss ratios:



From this graph, the ultimate loss ratios by year do not appear to be following a strong autocorrelation pattern, or other unexplained trends. If we had observed an increasing or decreasing pattern, then there could be a concern of bias introduced in our reserve estimate.

The following calculation shows the method of estimating reserves out to the 240 month evaluation point. As in the LDF method, this truncation point is used in order avoid undue reliance on a mechanical extrapolation formula.

The Cape Cod method works much like the more familiar Bornhuetter-Ferguson formula. Estimated reserves are calculated as a percent of the premium and the calculated expected loss ratio (ELR).

Accident Year	Onlevel Premium	Age at 12/31/2000	Average Age (x)	Growth Function	90.83% minus Growth Func	Premium x ELR	Estimated Reserves
		240	234	90.83%			
1991	10,000,000	120	114	77.76%	13.07%	5,977,659	781,218
1992	10,400,000	108	102	74.85%	15.98%	6,216,765	993,281
1993	10,800,000	96	90	71.29%	19.54%	6,455,872	1,261,416
1994	11,200,000	84	78	66.87%	23.96%	6,694,978	1,604,006
1995	11,600,000	72	66	61.31%	29.52%	6,934,085	2,046,646
1996	12,000,000	60	54	54.24%	36.59%	7,173,191	2,624,620
1997	12,400,000	48	42	45.17%	45.66%	7,412,297	3,384,400
1998	12,800,000	36	30	33.60%	57.22%	7,651,404	4,378,344
1999	13,200,000	24	18	19.46%	71.37%	7,890,510	5,631,298
2000	13,600,000	12	6	4.69%	86.13%	8,129,616	7,002,255
Total	118,000,000					70,536,377	29,707,484

For the variance calculation, we again begin with the process variance/mean ratio, which follows the chi-square formula. The sum of chi-square values is divided by 52 (55 data points minus 3 parameters), resulting in a σ^2 of 61,577. This turns out to be less than the 65,029 calculated for the LDF method because there we divided by 43 (55 data points minus 12 parameters).

The covariance matrix is estimated from the second derivative Information Matrix, and results in the following:

$$
\begin{array}{cccc}
 & \text{ELR} & \omega & \theta \\
\text{ELR} & 0.002421 & -0.002997 & 0.242396 \\
\omega & -0.002997 & 0.007853 & -0.401000 \\
\theta & 0.242396 & -0.401000 & 33.021994
\end{array}
$$

The standard deviation of our reserve estimate is calculated in the following table.

Accident Year	Reported Losses	Estimated Reserves	Process Std Dev	CV	Parameter Std Dev	CV	Total Std Dev	CV
1991	3,901,463	781,218	219,329	28.1%	158,913	20.3%	270,848	34.7%
1992	5,339,085	993,281	247,312	24.9%	192,103	19.3%	313,156	31.5%
1993	4,909,315	1,261,416	278,701	22.1%	229,523	18.2%	361,047	28.6%
1994	4,588,268	1,604,006	314,277	19.6%	270,790	16.9%	414,846	25.9%
1995	3,873,311	2,046,646	355,002	17.3%	314,629	15.4%	474,360	23.2%
1996	3,691,712	2,624,620	402,015	15.3%	358,200	13.6%	538,445	20.5%
1997	3,483,130	3,384,400	456,510	13.5%	396,353	11.7%	604,563	17.9%
1998	2,864,498	4,378,344	519,235	11.9%	421,934	9.6%	669,054	15.3%
1999	1,363,294	5,631,298	588,862	10.5%	430,873	7.7%	729,664	13.0%
2000	344,014	7,002,255	656,641	9.4%	439,441	6.3%	790,118	11.3%
Total	34,358,090	29,707,484	1,352,515	4.6%	3,143,967	10.6%	3,422,547	11.5%

In the earlier LDF example, the standard deviation on the overall reserve was 4,885,707 and this reduces to 3,422,547 when we switch to the Cape Cod method. The reduction is primarily seen in the more recent years 1999 and 2000, but is generally true for the full loss history. The reduction in the variance (the standard deviations squared) is even more extreme – <u>the overall variance in reserves is cut in half</u>.

This conclusion is at first surprising, since the two methods are very familiar to most actuaries. The difference is that we are making use of more information in the Cape Cod method, namely the onlevel premium by year, and this information allows us to make a significantly better estimate of the reserve.

4.3 Other Calculations Possible with this Model

Once the maximum likelihood calculations have been done, there are some other uses for the statistics besides the variance of the overall reserve. We will briefly look at three of these uses.

4.3.1 Variance of the Prospective Losses

Reserve reviews always focus on losses that have already occurred, but there is an intimate connection to the forecast of losses for the prospective period. The variability estimates from the Cape Cod method help us make this connection.

If the prospective period is estimated to include 14,000,000 in premium, we have a ready estimate of expected loss as 8,369,200 based on our 59.78% ELR. The process variance is calculated using the variance/mean multiplier 61,577, producing a CV of 8.6%.

The parameter variance is also readily calculated using the covariance matrix from the earlier calculation.

$$
\begin{array}{cc}
 & \begin{array}{ccc} ELR & \omega & \theta \end{array} \\
\begin{array}{c} ELR \\ \omega \\ \theta \end{array} &
\begin{pmatrix}
0.002421 & -0.002997 & 0.242396 \\
-0.002997 & 0.007853 & -0.401000 \\
0.242396 & -0.401000 & 33.021994
\end{pmatrix}
\end{array}
$$

The .002421 variance on the ELR translates to a standard deviation of 4.92% (by taking the square root) around our estimated ELR of 59.78%. Combined with the process variance, we have a total CV of 11.9%.

The CV from this estimate can then be compared to numbers produced by other prospective pricing tools.

4.3.2 *Calendar Year Development*

The stochastic reserving model can also be used to estimate development or payment for the next calendar year period beyond the latest diagonal. An example, using the LDF method is shown below.

Accident Year	Reported Losses	Age at 12/31/2000	Growth Function	Age at 12/31/2001	Growth Function	Estimated Ultimate	Est. 12 month Development
1991	3,901,463	120	77.24%	132	79.67%	5,050,867	122,450
1992	5,339,085	108	74.32%	120	77.24%	7,184,079	210,145
1993	4,909,315	96	70.75%	108	74.32%	6,939,399	247,928
1994	4,588,268	84	66.32%	96	70.75%	6,917,862	305,811
1995	3,873,311	72	60.78%	84	66.32%	6,372,348	353,146
1996	3,691,712	60	53.75%	72	60.78%	6,867,980	482,859
1997	3,483,130	48	44.77%	60	53.75%	7,780,515	699,093
1998	2,864,498	36	33.34%	48	44.77%	8,590,793	981,372
1999	1,363,294	24	19.38%	36	33.34%	7,033,659	981,996
2000	344,014	12	4.74%	24	19.38%	7,261,205	1,063,384
Total	34,358,090					69,998,708	5,448,182

The estimated development for the next 12-month calendar period is calculated by the difference in the growth functions at the two evaluation ages times the estimated ultimate losses. The standard deviation around this estimated development is:

Accident Year	Reported Losses	Est. 12 month Development	Process Std Dev	CV	Parameter Std Dev	CV	Total Std Dev	CV
1991	3,901,463	122,450	89,234	72.9%	24,632	20.1%	92,572	75.6%
1992	5,339,085	210,145	116,900	55.6%	37,767	18.0%	122,849	58.5%
1993	4,909,315	247,928	126,974	51.2%	42,716	17.2%	133,967	54.0%
1994	4,588,268	305,811	141,020	46.1%	50,260	16.4%	149,708	49.0%
1995	3,873,311	353,146	151,541	42.9%	57,208	16.2%	161,980	45.9%
1996	3,691,712	482,859	177,200	36.7%	74,987	15.5%	192,413	39.8%
1997	3,483,130	699,093	213,217	30.5%	106,043	15.2%	238,131	34.1%
1998	2,864,498	981,372	252,621	25.7%	158,978	16.2%	298,482	30.4%
1999	1,363,294	981,996	252,702	25.7%	225,920	23.0%	338,966	34.5%
2000	344,014	1,063,384	262,965	24.7%	480,861	45.2%	548,068	51.5%
Total	34,358,090	5,448,182	595,223	10.9%	635,609	11.7%	870,798	16.0%

A major reason for calculating the 12-month development is that the estimate is testable within a relatively short timeframe. If we project 5,448,182 of development, along with a standard deviation of 870,798, then one year later we can compare the actual development and see if it was within the forecast range.

4.3.3 Variability in Discounted Reserves

The mathematics for calculating the variability around discounted reserves follows directly from the payout pattern, model parameters and covariance matrix already calculated. The details are provided in Appendix C. This calculation is, of course, only appropriate if the analysis is being performed on paid data.

For the Cape Cod calculation of reserves, along with the 240 month truncation point, the discounted reserve using a 6.0% rate is provided below.

Accident Year	Estimated Reserves	Discounted Reserves	Process Std Dev	C.V.	Parameter Std Dev	C.V.	Total Std Dev	C.V.
1991	781,218	632,995	179,807	28.4%	125,961	19.9%	219,538	34.7%
1992	993,281	796,674	201,069	25.2%	149,689	18.8%	250,670	31.5%
1993	1,261,416	1,003,816	225,216	22.4%	175,899	17.5%	285,767	28.5%
1994	1,604,006	1,269,446	252,987	19.9%	204,084	16.1%	325,043	25.6%
1995	2,046,646	1,614,650	285,275	17.7%	232,952	14.4%	368,305	22.8%
1996	2,624,620	2,068,611	323,114	15.6%	259,904	12.6%	414,672	20.0%
1997	3,384,400	2,669,559	367,518	13.8%	280,605	10.5%	462,394	17.3%
1998	4,378,344	3,459,057	418,912	12.1%	289,876	8.4%	509,427	14.7%
1999	5,631,298	4,449,320	475,291	10.7%	286,857	6.4%	555,147	12.5%
2000	7,002,255	5,490,513	526,186	9.6%	284,582	5.2%	598,213	10.9%
Total	29,707,484	23,454,641	1,089,311	4.6%	2,198,224	9.4%	2,453,322	10.5%

From Section 4.2 above, we saw that the full-value reserve of 29,707,486 had a CV of 11.5%. The discounted reserve of 23,454,641 has a CV of 10.5%. The smaller CV for the discounted reserve is because the "tail" of the payout curve has the greatest parameter variance and also receives the deepest discount.

Section 5: Comments and Conclusion

5.1 Comments

Having worked through an example of stochastic reserving, a few practical comments are in order.

1) Abandon your triangles!

The maximum likelihood model works most logically from the tabular format of data as shown in tables 1.1 and 2.1. It is possible to first create the more familiar triangular format and then build the table, but there is no need for that intermediate step. All that is really needed is a consistent aggregation of losses evaluated at more than one date; we can skip the step of creating the triangle altogether.

2) The CV Goes with the Mean

The question of the use of the standard deviation or CV from the MLE is common. If we select a carried reserve other than the maximum likelihood estimate, then can we still use the CV from the model?

The short answer is "no". The estimate of the standard deviation in this model is very explicitly the standard deviation around the maximum likelihood estimate. If you do not trust the expected reserve from the MLE model, then there is even less reason to trust the standard deviation.

The more practical answer is an equivocal "yes". The final carried reserve is a selection, based on many factors including the use of a statistical model. No purely mechanical model should be the basis for setting the reserve, because it cannot take into account all of the characteristics of the underlying loss phenomenon. The standard deviation or CV around the selected reserve must therefore also be a selection, and a reasonable basis for that selection is the output of the MLE model.

The selection of a reserve range also needs to include consideration about changes in mix of business and the process of settling claims. These types of considerations might better be labeled "model variance", since by definition they are factors outside of the assumptions of the model.

3) Other Curve Forms

This paper has applied the method of maximum likelihood using growth curves that follow the Loglogistic and Weibull curve forms. These curves are useful in that they smoothly move from 0% to 100%, they often closely match the empirical data, and the first and second derivatives are calculable without the need for numerical approximations. However, the method in general is not limited to these forms and a larger library of curves can be investigated.

In this paper the Loglogistic and Weibull curves were applied to the average evaluation age, rather than the age from inception of the historical policy period. This was done for practical purposes, and is one way of improving the fit at immature ages. When evaluation ages fall within the period being developed (that is the period is not yet fully earned), then a further annualizing adjustment is needed. The formulas for this adjustment are given in Appendix B.

5.2 Conclusion

The method of maximum likelihood is a very useful technique for estimating both the expected development pattern and the variance around the estimated reserve. The use of the over-dispersed Poisson distribution is a convenient link to the LDF and Cape Cod estimates already common among reserving actuaries.

The chief result that we observe in working on practical examples is that the "parameter variance" component is generally larger than the "process variance" – most of the uncertainty in the estimated reserve is related to our inability to reliably estimate the expected reserve, not to random events. As such, our most pressing need is not for more sophisticated models, but for more complete data. Supplementing the standard loss development triangle with accident year exposure information is a good step in that direction.

Table 1.1
Original Triangle in Tabular Format

AY	From	To	Increment	Diag Age	AY Total
1991	0	12	357,848	120	3,901,463
1991	12	24	766,940	120	3,901,463
1991	24	36	610,542	120	3,901,463
1991	36	48	447,378	120	3,901,463
1991	48	60	562,888	120	3,901,463
1991	60	72	574,398	120	3,901,463
1991	72	84	146,342	120	3,901,463
1991	84	96	139,950	120	3,901,463
1991	96	108	227,229	120	3,901,463
1991	108	120	67,948	120	3,901,463
1992	0	12	352,118	108	5,339,085
1992	12	24	884,021	108	5,339,085
1992	24	36	933,894	108	5,339,085
1992	36	48	1,183,289	108	5,339,085
1992	48	60	445,745	108	5,339,085
1992	60	72	320,996	108	5,339,085
1992	72	84	527,804	108	5,339,085
1992	84	96	266,172	108	5,339,085
1992	96	108	425,046	108	5,339,085
1993	0	12	290,507	96	4,909,315
1993	12	24	1,001,799	96	4,909,315
1993	24	36	926,219	96	4,909,315
1993	36	48	1,016,654	96	4,909,315
1993	48	60	750,816	96	4,909,315
1993	60	72	146,923	96	4,909,315
1993	72	84	495,992	96	4,909,315
1993	84	96	280,405	96	4,909,315
1994	0	12	310,608	84	4,588,268
1994	12	24	1,108,250	84	4,588,268
1994	24	36	776,189	84	4,588,268
1994	36	48	1,562,400	84	4,588,268
1994	48	60	272,482	84	4,588,268
1994	60	72	352,053	84	4,588,268
1994	72	84	206,286	84	4,588,268
1995	0	12	443,160	72	3,873,311
1995	12	24	693,190	72	3,873,311
1995	24	36	991,983	72	3,873,311
1995	36	48	769,488	72	3,873,311
1995	48	60	504,851	72	3,873,311
1995	60	72	470,639	72	3,873,311
1996	0	12	396,132	60	3,691,712
1996	12	24	937,085	60	3,691,712
1996	24	36	847,498	60	3,691,712
1996	36	48	805,037	60	3,691,712
1996	48	60	705,960	60	3,691,712
1997	0	12	440,832	48	3,483,130
1997	12	24	847,631	48	3,483,130
1997	24	36	1,131,398	48	3,483,130
1997	36	48	1,063,269	48	3,483,130
1998	0	12	359,480	36	2,864,498
1998	12	24	1,061,648	36	2,864,498
1998	24	36	1,443,370	36	2,864,498
1999	0	12	376,686	24	1,363,294
1999	12	24	986,608	24	1,363,294
2000	0	12	344,014	12	344,014

Table 1.2
Triangle Collapsed for Latest Three Diagonals

AY	From	To	Increment	Diag Age	AY Total
1991	0	96	3,606,286	120	3,901,463
1991	96	108	227,229	120	3,901,463
1991	108	120	67,948	120	3,901,463
1992	0	84	4,647,867	108	5,339,085
1992	84	96	266,172	108	5,339,085
1992	96	108	425,046	108	5,339,085
1993	0	72	4,132,918	96	4,909,315
1993	72	84	495,992	96	4,909,315
1993	84	96	280,405	96	4,909,315
1994	0	60	4,029,929	84	4,588,268
1994	60	72	352,053	84	4,588,268
1994	72	84	206,286	84	4,588,268
1995	0	48	2,897,821	72	3,873,311
1995	48	60	504,851	72	3,873,311
1995	60	72	470,639	72	3,873,311
1996	0	36	2,180,715	60	3,691,712
1996	36	48	805,037	60	3,691,712
1996	48	60	705,960	60	3,691,712
1997	0	24	1,288,463	48	3,483,130
1997	24	36	1,131,398	48	3,483,130
1997	36	48	1,063,269	48	3,483,130
1998	0	12	359,480	36	2,864,498
1998	12	24	1,061,648	36	2,864,498
1998	24	36	1,443,370	36	2,864,498
1999	0	12	376,686	24	1,363,294
1999	12	24	986,608	24	1,363,294
2000	0	12	344,014	12	344,014

Table 1.3
Latest Diagonal Representing only 9 Months of Development

AY	From	To	Increment	Diag Age	AY Total
1991	0	12	357,848	117	3,901,463
1991	12	24	766,940	117	3,901,463
1991	24	36	610,542	117	3,901,463
1991	36	48	447,378	117	3,901,463
1991	48	60	562,888	117	3,901,463
1991	60	72	574,398	117	3,901,463
1991	72	84	146,342	117	3,901,463
1991	84	96	139,950	117	3,901,463
1991	96	108	227,229	117	3,901,463
1991	108	117	67,948	117	3,901,463
1992	0	12	352,118	105	5,339,085
1992	12	24	884,021	105	5,339,085
1992	24	36	933,894	105	5,339,085
1992	36	48	1,183,289	105	5,339,085
1992	48	60	445,745	105	5,339,085
1992	60	72	320,996	105	5,339,085
1992	72	84	527,804	105	5,339,085
1992	84	96	266,172	105	5,339,085
1992	96	105	425,046	105	5,339,085
1993	0	12	290,507	93	4,909,315
1993	12	24	1,001,799	93	4,909,315
1993	24	36	926,219	93	4,909,315
1993	36	48	1,016,654	93	4,909,315
1993	48	60	750,816	93	4,909,315
1993	60	72	146,923	93	4,909,315
1993	72	84	495,992	93	4,909,315
1993	84	93	280,405	93	4,909,315
1994	0	12	310,608	81	4,588,268
1994	12	24	1,108,250	81	4,588,268
1994	24	36	776,189	81	4,588,268
1994	36	48	1,562,400	81	4,588,268
1994	48	60	272,482	81	4,588,268
1994	60	72	352,053	81	4,588,268
1994	72	81	206,286	81	4,588,268
1995	0	12	443,160	69	3,873,311
1995	12	24	693,190	69	3,873,311
1995	24	36	991,983	69	3,873,311
1995	36	48	769,488	69	3,873,311
1995	48	60	504,851	69	3,873,311
1995	60	69	470,639	69	3,873,311
1996	0	12	396,132	57	3,691,712
1996	12	24	937,085	57	3,691,712
1996	24	36	847,498	57	3,691,712
1996	36	48	805,037	57	3,691,712
1996	48	57	705,960	57	3,691,712
1997	0	12	440,832	45	3,483,130
1997	12	24	847,631	45	3,483,130
1997	24	36	1,131,398	45	3,483,130
1997	36	45	1,063,269	45	3,483,130
1998	0	12	359,480	33	2,864,498
1998	12	24	1,061,648	33	2,864,498
1998	24	33	1,443,370	33	2,864,498
1999	0	12	376,686	21	1,363,294
1999	12	21	986,608	21	1,363,294
2000	0	9	344,014	9	344,014

Table 1.4
Original Triangle along with Fitted Values – LDF Method

AY	From	To	Increment	Diag Age	AY Total	Est. ULT	Fitted	MLE Term	Chi-Square
1991	0	12	357,848	120	3,901,463	5,050,868	239,295	4,192,814	58,734
1991	12	24	766,940	120	3,901,463	5,050,868	739,686	9,624,727	1,004
1991	24	36	610,542	120	3,901,463	5,050,868	705,171	7,516,507	12,698
1991	36	48	447,378	120	3,901,463	5,050,868	576,987	5,357,739	29,114
1991	48	60	562,888	120	3,901,463	5,050,868	453,829	6,878,055	26,208
1991	60	72	574,398	120	3,901,463	5,050,868	355,106	6,985,799	135,422
1991	72	84	146,342	120	3,901,463	5,050,868	279,911	1,555,543	63,737
1991	84	96	139,950	120	3,901,463	5,050,868	223,278	1,500,370	31,098
1991	96	108	227,229	120	3,901,463	5,050,868	180,455	2,569,751	12,124
1991	108	120	67,948	120	3,901,463	5,050,868	147,745	661,056	43,099
1992	0	12	352,118	108	5,339,085	7,184,081	340,360	4,144,834	406
1992	12	24	884,021	108	5,339,085	7,184,081	1,052,089	11,206,001	26,848
1992	24	36	933,894	108	5,339,085	7,184,081	1,002,997	11,902,020	4,761
1992	36	48	1,183,289	108	5,339,085	7,184,081	820,675	15,293,216	160,220
1992	48	60	445,745	108	5,339,085	7,184,081	645,502	5,317,578	61,817
1992	60	72	320,996	108	5,339,085	7,184,081	505,083	3,710,390	67,094
1992	72	84	527,804	108	5,339,085	7,184,081	398,131	6,407,657	42,235
1992	84	96	266,172	108	5,339,085	7,184,081	317,579	3,054,416	8,321
1992	96	108	425,046	108	5,339,085	7,184,081	256,670	5,037,510	110,456
1993	0	12	290,507	96	4,909,315	6,939,401	328,768	3,361,574	4,453
1993	12	24	1,001,799	96	4,909,315	6,939,401	1,016,256	12,840,263	206
1993	24	36	926,219	96	4,909,315	6,939,401	968,836	11,798,028	1,875
1993	36	48	1,016,654	96	4,909,315	6,939,401	792,724	13,016,722	63,256
1993	48	60	750,816	96	4,909,315	6,939,401	623,517	9,394,719	25,990
1993	60	72	146,923	96	4,909,315	6,939,401	487,881	1,436,491	238,280
1993	72	84	495,992	96	4,909,315	6,939,401	384,571	5,993,828	32,282
1993	84	96	280,405	96	4,909,315	6,939,401	306,763	3,235,826	2,265
1994	0	12	310,608	84	4,588,268	6,917,864	327,748	3,616,974	896
1994	12	24	1,108,250	84	4,588,268	6,917,864	1,013,102	14,312,364	8,936
1994	24	36	776,189	84	4,588,268	6,917,864	965,829	9,730,631	37,236
1994	36	48	1,562,400	84	4,588,268	6,917,864	790,264	20,427,319	754,424
1994	48	60	272,482	84	4,588,268	6,917,864	621,582	3,013,334	196,065
1994	60	72	352,053	84	4,588,268	6,917,864	486,366	4,123,668	37,092
1994	72	84	206,286	84	4,588,268	6,917,864	383,377	2,268,795	81,803
1995	0	12	443,160	72	3,873,311	6,372,350	301,903	5,289,828	66,093
1995	12	24	693,190	72	3,873,311	6,372,350	933,213	8,595,646	61,734
1995	24	36	991,983	72	3,873,311	6,372,350	889,668	12,699,114	11,767
1995	36	48	769,488	72	3,873,311	6,372,350	727,947	9,658,589	2,371
1995	48	60	504,851	72	3,873,311	6,372,350	572,566	6,120,690	8,008
1995	60	72	470,639	72	3,873,311	6,372,350	448,014	5,676,214	1,143
1996	0	12	396,132	60	3,691,712	6,867,982	325,384	4,702,625	15,382
1996	12	24	937,085	60	3,691,712	6,867,982	1,005,797	11,945,927	4,694
1996	24	36	847,498	60	3,691,712	6,867,982	958,865	10,714,153	12,935
1996	36	48	805,037	60	3,691,712	6,867,982	784,566	10,142,109	534
1996	48	60	705,960	60	3,691,712	6,867,982	617,100	8,795,314	12,796
1997	0	12	440,832	48	3,483,130	7,780,518	368,618	5,281,753	14,147
1997	12	24	847,631	48	3,483,130	7,780,518	1,139,436	10,681,663	74,730
1997	24	36	1,131,398	48	3,483,130	7,780,518	1,086,268	14,638,194	1,875
1997	36	48	1,063,269	48	3,483,130	7,780,518	888,809	13,675,465	34,244
1998	0	12	359,480	36	2,864,498	8,590,795	407,006	4,236,247	5,550
1998	12	24	1,061,648	36	2,864,498	8,590,795	1,258,098	13,652,867	30,675
1998	24	36	1,443,370	36	2,864,498	8,590,795	1,199,393	19,003,928	49,629
1999	0	12	376,686	24	1,363,294	7,033,660	333,234	4,456,931	5,666
1999	12	24	986,608	24	1,363,294	7,033,660	1,030,060	12,629,654	1,833
2000	0	12	344,014	12	344,014	7,261,202	344,014	4,041,627	0
			34,358,090				34,358,090		2,796,260

Table 2.1
Original Triangle along with Fitted Values – Cape Cod Method

AY	From	To	Increment	Diag Age	Premium	Est. ULT	Fitted	MLE Term	Chi-Square
1991	0	12	357,848	120	10,000,000	5,977,659	280,569	4,208,482	21,285
1991	12	24	766,940	120	10,000,000	5,977,659	882,582	9,617,292	15,152
1991	24	36	610,542	120	10,000,000	5,977,659	845,554	7,486,969	65,319
1991	36	48	447,378	120	10,000,000	5,977,659	691,227	5,324,318	86,024
1991	48	60	562,888	120	10,000,000	5,977,659	542,171	6,889,829	792
1991	60	72	574,398	120	10,000,000	5,977,659	422,833	7,018,339	54,329
1991	72	84	146,342	120	10,000,000	5,977,659	332,202	1,528,317	103,985
1991	84	96	139,950	120	10,000,000	5,977,659	264,171	1,483,014	58,412
1991	96	108	227,229	120	10,000,000	5,977,659	212,900	2,574,877	964
1991	108	120	67,948	120	10,000,000	5,977,659	173,860	646,001	64,519
1992	0	12	352,118	108	10,400,000	6,216,765	291,792	4,139,189	12,472
1992	12	24	884,021	108	10,400,000	6,216,765	917,885	11,219,571	1,249
1992	24	36	933,894	108	10,400,000	6,216,765	879,376	11,902,801	3,380
1992	36	48	1,183,289	108	10,400,000	6,216,765	718,876	15,238,302	300,023
1992	48	60	445,745	108	10,400,000	6,216,765	563,858	5,338,946	24,742
1992	60	72	320,996	108	10,400,000	6,216,765	439,746	3,731,261	32,068
1992	72	84	527,804	108	10,400,000	6,216,765	345,490	6,385,446	96,207
1992	84	96	266,172	108	10,400,000	6,216,765	274,738	3,058,687	267
1992	96	108	425,046	108	10,400,000	6,216,765	221,416	5,009,964	187,273
1993	0	12	290,507	96	10,800,000	6,455,872	303,015	3,363,630	516
1993	12	24	1,001,799	96	10,800,000	6,455,872	953,188	12,839,147	2,479
1993	24	36	926,219	96	10,800,000	6,455,872	913,198	11,798,887	186
1993	36	48	1,016,654	96	10,800,000	6,455,872	746,525	13,001,875	97,746
1993	48	60	750,816	96	10,800,000	6,455,872	585,545	9,385,515	46,648
1993	60	72	146,923	96	10,800,000	6,455,872	456,660	1,457,996	210,084
1993	72	84	495,992	96	10,800,000	6,455,872	358,778	5,985,187	52,477
1993	84	96	280,405	96	10,800,000	6,455,872	285,305	3,236,950	84
1994	0	12	310,608	84	11,200,000	6,694,978	314,238	3,617,409	42
1994	12	24	1,108,250	84	11,200,000	6,694,978	988,491	14,309,720	14,509
1994	24	36	776,189	84	11,200,000	6,694,978	947,020	9,734,175	30,816
1994	36	48	1,562,400	84	11,200,000	6,694,978	774,174	20,411,270	802,533
1994	48	60	272,482	84	11,200,000	6,694,978	607,232	3,021,320	184,538
1994	60	72	352,053	84	11,200,000	6,694,978	473,573	4,127,077	31,182
1994	72	84	206,286	84	11,200,000	6,694,978	372,066	2,273,929	73,866
1995	0	12	443,160	72	11,600,000	6,934,085	325,460	5,299,568	42,565
1995	12	24	693,190	72	11,600,000	6,934,085	1,023,795	8,569,280	106,759
1995	24	36	991,983	72	11,600,000	6,934,085	980,842	12,704,721	127
1995	36	48	769,488	72	11,600,000	6,934,085	801,823	9,659,092	1,304
1995	48	60	504,851	72	11,600,000	6,934,085	628,919	6,111,729	24,475
1995	60	72	470,639	72	11,600,000	6,934,085	490,486	5,676,368	803
1996	0	12	396,132	60	12,000,000	7,173,191	336,683	4,704,848	10,497
1996	12	24	937,085	60	12,000,000	7,173,191	1,059,098	11,941,015	14,056
1996	24	36	847,498	60	12,000,000	7,173,191	1,014,664	10,706,291	27,541
1996	36	48	805,037	60	12,000,000	7,173,191	829,472	10,142,011	720
1996	48	60	705,960	60	12,000,000	7,173,191	650,606	8,799,134	4,710
1997	0	12	440,832	48	12,400,000	7,412,297	347,906	5,276,973	24,821
1997	12	24	847,631	48	12,400,000	7,412,297	1,094,401	10,692,516	55,643
1997	24	36	1,131,398	48	12,400,000	7,412,297	1,048,487	14,635,924	6,556
1997	36	48	1,063,269	48	12,400,000	7,412,297	857,121	13,668,552	49,581
1998	0	12	359,480	36	12,800,000	7,651,404	359,129	4,239,137	0
1998	12	24	1,061,648	36	12,800,000	7,651,404	1,129,704	13,666,979	4,100
1998	24	36	1,443,370	36	12,800,000	7,651,404	1,082,309	18,972,750	120,451
1999	0	12	376,686	24	13,200,000	7,890,510	370,351	4,459,595	108
1999	12	24	986,608	24	13,200,000	7,890,510	1,165,008	12,616,168	27,319
2000	0	12	344,014	12	13,600,000	8,129,616	381,574	4,039,715	3,697
			34,358,090				34,358,090		3,202,001

References:

[1] England & Verrall; *A Flexible Framework for Stochastic Claims Reserving*; Proceedings of the Casualty Actuarial Society, 2001.

[2] Halliwell, Leigh; *Loss Prediction by Generalized Least Squares*; Proceedings of the Casualty Actuarial Society, 1996.

[3] Klugman, Panjer & Wilmot; Loss Models: From Data to Decisions; John Wiley & Sons, 1998.
See mainly pages 60-68, for discussion of maximum likelihood.

[4] Kreps, Rodney; *Parameter Uncertainty in (Log)Normal Distributions*; Proceedings of the Casualty Actuarial Society, 1997.

[5] Mack, Thomas; *A Simple Parametric Model for Rating Automobile Insurance or Estimating IBNR Claims Reserve;* 1991 ASTIN Bulletin 21 (1), 93-109.

[6] Mack, Thomas; *Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates;* 1993 ASTIN Bulletin 23 (2), 213-225.

[7] McCullagh, P. & Nelder, J.A.; Generalized Linear Models; Chapman & Hall/CRC, second edition 1989.
See Chapter 6 on Log-linear Models, and the over-dispersed Poisson.

[8] Sherman, Richard E.; *Extrapolating, Smoothing, and Interpolating Development Factors*; Proceedings of the Casualty Actuarial Society, 1984.

[9] Weissner, Edward W.; *Estimation of the Distribution of Report Lags by the Method of Maximum Likelihood*; Proceedings of the Casualty Actuarial Society, 1978.

[10] Zehnwirth, Ben; *Probabilistic Development Factor Models with Applications to Loss Reserve Variability, Prediction Intervals and Risk Based Capital.* CAS Forum Spring 1994, Volume Two.

Appendix A: Derivatives of the Loglikelihood Function

The loglikelihood function for the over-dispersed Poisson is proportional to

$$\ell = \sum_i c_i \cdot \ln(\mu_i) - \mu_i$$

$$\text{where } \mu_{i,t} = ELR \cdot P_i \cdot [G(x_t \mid \omega,\theta) - G(x_{t-1} \mid \omega,\theta)]$$

as described in section 2.2 of this paper. The derivatives below are then used to complete the Information Matrix needed in the parameter variance calculation.

The derivatives of the exact loglikelihood function would require dividing all of these numbers by the constant scale factor σ^2, but it is easier to omit that here and apply it to the final covariance matrix at the end.

$$\frac{\partial^2 \ell}{\partial ELR^2} = \sum_{i,t} \left(\frac{-c_{i,t}}{ELR^2} \right)$$

$$\frac{\partial^2 \ell}{\partial ELR \partial \omega} = -\sum_{i,t} P_i \cdot \left[\frac{\partial G(x_t)}{\partial \omega} - \frac{\partial G(x_{t-1})}{\partial \omega} \right]$$

$$\frac{\partial^2 \ell}{\partial ELR \partial \theta} = -\sum_{i,t} P_i \cdot \left[\frac{\partial G(x_t)}{\partial \theta} - \frac{\partial G(x_{t-1})}{\partial \theta} \right]$$

$$\frac{\partial \ell}{\partial \omega} = \sum_{i,t} \left\{ \left[\frac{c_{i,t}}{G(x_t) - G(x_{t-1})} - ELR \cdot P_i \right] \cdot \left[\frac{\partial G(x_t)}{\partial \omega} - \frac{\partial G(x_{t-1})}{\partial \omega} \right] \right\}$$

$$\frac{\partial^2 \ell}{\partial \omega^2} = \sum_{i,t} \left\{ \left[\frac{-c_{i,t}}{(G(x_t) - G(x_{t-1}))^2} \right] \cdot \left[\frac{\partial G(x_t)}{\partial \omega} - \frac{\partial G(x_{t-1})}{\partial \omega} \right]^2 + \left[\frac{c_{i,t}}{G(x_t) - G(x_{t-1})} - ELR \cdot P_i \right] \cdot \left[\frac{\partial^2 G(x_t)}{\partial \omega^2} - \frac{\partial^2 G(x_{t-1})}{\partial \omega^2} \right] \right\}$$

$$\frac{\partial^2 \ell}{\partial\omega\,\partial\theta} = \sum_{i,t}\left\{\left[\frac{-c_{i,t}}{(G(x_t)-G(x_{t-1}))^2}\right]\cdot\left[\frac{\partial G(x_t)}{\partial\omega}-\frac{\partial G(x_{t-1})}{\partial\omega}\right]\left[\frac{\partial G(x_t)}{\partial\theta}-\frac{\partial G(x_{t-1})}{\partial\theta}\right]+\right.$$

$$\left.\left[\frac{c_{i,t}}{G(x_t)-G(x_{t-1})}-ELR\cdot P_i\right]\cdot\left[\frac{\partial^2 G(x_t)}{\partial\omega\,\partial\theta}-\frac{\partial^2 G(x_{t-1})}{\partial\omega\partial\theta}\right]\right\}$$

$$\frac{\partial \ell}{\partial\theta} = \sum_{i,t}\left\{\left[\frac{c_{i,t}}{G(x_t)-G(x_{t-1})}-ELR\cdot P_i\right]\cdot\left[\frac{\partial G(x_t)}{\partial\theta}-\frac{\partial G(x_{t-1})}{\partial\theta}\right]\right\}$$

$$\frac{\partial^2 \ell}{\partial\theta^2} = \sum_{i,t}\left\{\left[\frac{-c_{i,t}}{(G(x_t)-G(x_{t-1}))^2}\right]\cdot\left[\frac{\partial G(x_t)}{\partial\theta}-\frac{\partial G(x_{t-1})}{\partial\theta}\right]^2+\right.$$

$$\left.\left[\frac{c_{i,t}}{G(x_t)-G(x_{t-1})}-ELR\cdot P_i\right]\cdot\left[\frac{\partial^2 G(x_t)}{\partial\theta^2}-\frac{\partial^2 G(x_{t-1})}{\partial\theta^2}\right]\right\}$$

For the LDF Method, these same formulas apply but replacing:

$ELR \rightarrow ULT_i$ and $P_i \rightarrow 1$.

Weibull Distribution

$$G(x) = F(x) = 1 - \exp\left[-(x/\theta)^{\omega}\right]$$

$$f(x) = \frac{\omega}{x}\cdot\left(\frac{x}{\theta}\right)^{\omega}\cdot\exp\left[-(x/\theta)^{\omega}\right]$$

$$E[x^k] = \theta^k\cdot\Gamma(1+k/\omega)$$

θ is approximately the 63.2%-tile $= 1-\exp[-1]$, $LDF_{\theta} \approx 1.582$

$$\frac{\partial G(x)}{\partial\omega} = \exp\left[-\left(\tfrac{x}{\theta}\right)^{\omega}\right]\cdot\left(\frac{x}{\theta}\right)^{\omega}\cdot\ln\left(\tfrac{x}{\theta}\right)$$

$$\frac{\partial G(x)}{\partial\theta} = \exp\left[-\left(\tfrac{x}{\theta}\right)^{\omega}\right]\cdot\left(\frac{x}{\theta}\right)^{\omega}\cdot\left(\frac{-\omega}{\theta}\right)$$

$$\frac{\partial^2 G(x)}{\partial\omega^2} = \exp\left[-\left(\tfrac{x}{\theta}\right)^{\omega}\right]\cdot\left(\frac{x}{\theta}\right)^{\omega}\cdot\ln\left(\tfrac{x}{\theta}\right)^2\cdot\left[1-\left(\frac{x}{\theta}\right)^{\omega}\right]$$

$$\frac{\partial^2 G(x)}{\partial\omega\partial\theta} = \exp\left[-\left(\tfrac{x}{\theta}\right)^{\omega}\right]\cdot\left(\frac{x}{\theta}\right)^{\omega}\cdot\left(\frac{-1}{\theta}\right)\cdot\left\{1+\omega\cdot\ln\left(\tfrac{x}{\theta}\right)\cdot\left[1-\left(\frac{x}{\theta}\right)^{\omega}\right]\right\}$$

$$\frac{\partial^2 G(x)}{\partial\theta^2} = \exp\left[-\left(\tfrac{x}{\theta}\right)^{\omega}\right]\cdot\left(\frac{x}{\theta}\right)^{\omega}\cdot\left(\frac{\omega}{\theta^2}\right)\cdot\left\{1+\omega\cdot\left[1-\left(\frac{x}{\theta}\right)^{\omega}\right]\right\}$$

Loglogistic Distribution (for "inverse power" LDFs)

$$G(x) = F(x) = \frac{x^\omega}{x^\omega + \theta^\omega} = 1 - \left(\frac{1}{1+(x/\theta)^\omega}\right)$$

$$f(x) = \frac{\omega}{x} \cdot \left(\frac{x^\omega}{x^\omega + \theta^\omega}\right) \cdot \left(\frac{\theta^\omega}{x^\omega + \theta^\omega}\right)$$

$$E[x^k] = \theta^k \cdot \Gamma(1 + k/\omega) \cdot \Gamma(1 - k/\omega)$$

θ is the median of the distribution $LDF_\theta = 2.000$

$$\frac{\partial G(x)}{\partial \omega} = \left(\frac{x^\omega}{x^\omega + \theta^\omega}\right) \cdot \left(\frac{\theta^\omega}{x^\omega + \theta^\omega}\right) \cdot \ln\left(\tfrac{x}{\theta}\right)$$

$$\frac{\partial G(x)}{\partial \theta} = \left(\frac{x^\omega}{x^\omega + \theta^\omega}\right) \cdot \left(\frac{\theta^\omega}{x^\omega + \theta^\omega}\right) \cdot \left(\frac{-\omega}{\theta}\right)$$

$$\frac{\partial^2 G(x)}{\partial \omega^2} = \left(\frac{x^\omega}{x^\omega + \theta^\omega}\right) \cdot \left(\frac{\theta^\omega}{x^\omega + \theta^\omega}\right) \cdot \ln\left(\tfrac{x}{\theta}\right)^2 \cdot \left[1 - 2 \cdot \left(\frac{x^\omega}{x^\omega + \theta^\omega}\right)\right]$$

$$\frac{\partial^2 G(x)}{\partial \omega \partial \theta} = \left(\frac{x^\omega}{x^\omega + \theta^\omega}\right) \cdot \left(\frac{\theta^\omega}{x^\omega + \theta^\omega}\right) \cdot \left(\frac{-1}{\theta}\right) \cdot \left\{1 + \omega \cdot \ln\left(\tfrac{x}{\theta}\right) \cdot \left[1 - 2 \cdot \left(\frac{x^\omega}{x^\omega + \theta^\omega}\right)\right]\right\}$$

$$\frac{\partial^2 G(x)}{\partial \theta^2} = \left(\frac{x^\omega}{x^\omega + \theta^\omega}\right) \cdot \left(\frac{\theta^\omega}{x^\omega + \theta^\omega}\right) \cdot \left(\frac{\omega}{\theta^2}\right) \cdot \left\{1 + \omega \cdot \left[1 - 2 \cdot \left(\frac{x^\omega}{x^\omega + \theta^\omega}\right)\right]\right\}$$

Appendix B: Adjustments for Different Exposure Periods

The percent of ultimate curve is assumed to be a function of the average accident date of the period being developed to ultimate.

$$G^*(x \mid \omega,\theta) \quad = \text{cumulative percent of ultimate as of average date } x$$

Further, we will assume that this is the percent of ultimate for the portion of the period that has already been earned. For example, if we are 9 months into an accident year, then the quantity $G^*(4.5 \mid \omega,\theta)$ represents the cumulative percent of ultimate of the 9-month period only. The loss development factor $LDF_9^* = 1/G^*(4.5 \mid \omega,\theta)$ is the adjustment needed to calculate the ultimate loss dollars for the 9-month period (before annualizing).

In order to estimate the cumulative percent of ultimate for the full accident year, we also need to multiply by a scaling factor representing the portion of the accident year that has been earned.

The AY cumulative percent of ultimate as of 9 months is

$$G_{AY}(9 \mid \omega,\theta) \quad = \quad \left(\frac{9}{12}\right) \cdot G^*(4.5 \mid \omega,\theta)$$

We find therefore that we need to make two calculations:

1) Calculate the percent of the period that is exposed; *Expos(t)*
2) Calculate the average accident date given the age from inception t; *AvgAge(t)*

These functions can be easily calculated for accident year or policy year periods.

1) Calculate the percent of the period that is exposed: $Expos(t)$

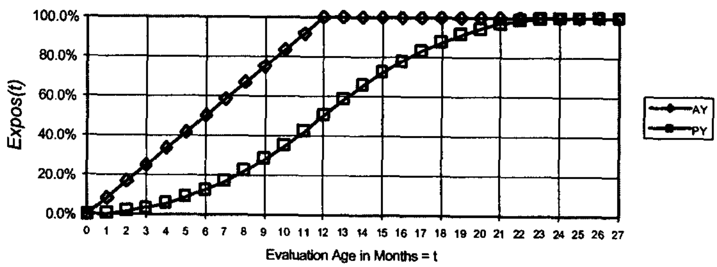
For accident years (AY):

$$Expos(t) = \begin{cases} t/12 & t \le 12 \\ 1 & t > 12 \end{cases} \quad \text{or}$$

For policy years (PY):

$$Expos(t) = \begin{cases} \frac{1}{2}\cdot(t/12)^2 & t \le 12 \\ 1-\frac{1}{2}\cdot\max(2-t/12,0)^2 & t > 12 \end{cases}$$

Cumulative Percent of Exposure Expos(t)



2) Calculate the average accident date of the period that is earned: $AvgAge(t)$

For accident years (AY):

$$AvgAge(t) = \begin{cases} t/2 & t \le 12 \\ t-6 & t > 12 \end{cases} \qquad \text{or} \qquad AvgAge(t) = \max(t-6, t/2)$$

For policy years (PY):

$$AvgAge(t) = \begin{cases} t/3 & t \le 12 \\ \dfrac{(t-12)+\frac{1}{3}\cdot(24-t)\cdot\left(1-Expos(t)\right)}{Expos(t)} & t > 12 \end{cases}$$

The final cumulative percent of ultimate curve, including annualization, is given by:

$$\boxed{G_{AY\,or\,PY}(t \mid \omega,\theta) = Expos(t)\cdot G^{*}\left(AvgAge(t) \mid \omega,\theta\right)}$$

Appendix C: Variance in Discounted Reserves

The maximum likelihood estimation model allows for the estimation of variance of discounted reserves as well as the variance of the full-value reserves. These calculations are a bit more tedious, and so are given just in this appendix.

Calculation of Discounted Reserve

We begin by recalling that the reserve is estimated as a sum of portions of all the historical accident years, and is calculated as:

$$\text{Reserve:} \qquad R \;=\; \sum_{AY} \mu_{AY;x,y} \;=\; \sum_{AY} ULT_{AY} \cdot \left(G(y) - G(x)\right)$$

This expression can be expanded as the sum of individual increments.

$$R \;=\; \sum_{AY} \sum_{k=1}^{y-x} ULT_{AY} \cdot \left(G(x+k) - G(x+k-1)\right)$$

To be even more precise, we could write this as a continuous function.

$$R \;=\; \sum_{AY} ULT_{AY} \cdot \int_x^y g(t)\,dt \qquad \text{where } \; g(t) = \frac{\partial G(t)}{\partial t}$$

The value of the discounted reserve R_d would then be written as follows.

$$R_d \;=\; \sum_{AY} ULT_{AY} \cdot \int_x^y v^{t-x} \cdot g(t)\,dt \qquad \text{where } \; v = \frac{1}{1+i}$$

For purposes of this paper, we will assume that the discount rate i is constant. There is also some debate as to what this rate should be (cost of capital?, market yield?), but we will avoid that discussion here.

An interesting note on this expression is seen in the case of $x=0$ and $y=\infty$, in which the form of the discounted loss at time zero is directly related to the moment generating function of the growth curve.

$$\int_0^\infty v^t \cdot g(y)\,dt \;=\; \int_0^\infty e^{-t \cdot \ln(1+i)} \cdot g(t)\,dt \;=\; MGF\left(-\ln(1+i)\right)$$

Unfortunately, for the Loglogistic and Weibull growth curves, the moment generating function is intractable and so does not simplify our calculation. For practical purposes we will use the incremental approximation instead.

$$R_d \approx \sum_{AY}\sum_{k=1}^{y-x} ULT_{AY} \cdot v^{k-1/2} \cdot \left(G(x+k) - G(x+k-1)\right)$$

The variance can then be calculated for the discounted reserve in two pieces: the process variance and the parameter variance.

Process Variance

The process variance component is actually trivial to calculate. We already know that the variance of the full value reserve is estimated by multiplying by the scale factor σ^2. We then need to recall that the variance for some random variable times a constant is given by $Var(v^k \cdot R) = v^{2k} \cdot Var(R)$.

The process variance of the discounted reserve is therefore:

$$Var(R_d) \approx \sigma^2 \cdot \sum_{AY}\sum_{k=1}^{y-x} ULT_{AY} \cdot v^{2k-1} \cdot \left(G(x+k) - G(x+k-1)\right)$$

Parameter Variance

The parameter variance again makes use of the covariance matrix of the model parameters Σ. The formula is then given below.

$$Var(E[R_d]) = (\partial R_d)' \cdot \Sigma \cdot (\partial R_d)$$

where

$$\partial R_d = \left\langle \frac{\partial R_d}{\partial ELR}, \frac{\partial R_d}{\partial \omega}, \frac{\partial R_d}{\partial \theta} \right\rangle \quad \text{for the Cape Cod method}$$

or

$$\partial R_d = \left\langle \left\{ \frac{\partial R_d}{\partial ULT_{AY}} \right\}_{AY=1}^{n}, \frac{\partial R_d}{\partial \omega}, \frac{\partial R_d}{\partial \theta} \right\rangle \quad \text{for the LDF method}$$

In order to calculate the derivatives of the discounted reserves, we make use of the same mathematical expressions as for the full value reserves. That is,

$$\frac{\partial R}{\partial \omega} = \sum_{AY,x} \frac{\partial \mu_{AY,x}}{\partial \omega} \qquad \text{becomes} \qquad \frac{\partial R_d}{\partial \omega} = \sum_{AY,x} v_{AY,x} \cdot \frac{\partial \mu_{AY,x}}{\partial \omega}$$

The calculation is similar to the variance calculation for the full value reserve, but now it is expanded for each increment so that the time dimension is included. The complexity of the calculations does not change, but the number of times they are performed greatly increases.

The combination of the process and parameter variances is simple addition, the same as for the full value reserves, since we make the assumption that the two sources of variance are independent.