

Loss Reserving with Random Selection

Abstract

This paper presents a random selection method with the Monte Carlo simulation technique in the estimation of loss reserves. The future loss development factors are randomly selected from a weighted empirical distribution of observed loss-development factors. This nonparametric approach provides an estimate of the distribution of total loss reserve. By assigning proper weights, the mean of this distribution is statistically equivalent to the result from the traditional Chain-Ladder method. The variance of total loss reserve can also be approximated through this approach. In general, the proposed method is very flexible and can be easily extended to many circumstances, including the Bornheutter-Ferguson (BF) method (Bornheutter and Ferguson, 1972). The results are further enhanced by implementing the simulation scheme with smoothing techniques.

Keywords: Chain-Ladder method; Random selection; Loss reserving.

1 Introduction

The task of estimating the amount of current liability associated with the future contingent loss event is known as loss reserving in the insurance industry. The insurer is liable on the date of occurrence of claim. Typically, the ultimate value of claim will not be paid and known until some time later. For practical purposes, it is very important that an insurance company accrues a reserve for its future payments. When the level of reserve is inadequate, it could lead to insolvency.

There are many loss reserving methods, some deterministic and some stochastic. Usually, the actuary will choose several methods for the purpose of determining a reserve estimate. For a comprehensive review of loss reserving methods, the reader is referred to Bornheutter and Ferguson (1972), Finger (1976), Taylor (2000), Foundations of Casualty Actuarial Science (2001), Brown and Gottlieb (2001), and England and Verrall (2002).

Recently, it has become more evident that actuaries also need some measures of dispersion for a loss reserve estimation. It can be used to develop a confidence interval for a prudential margin check. Many researchers have developed different methods to describe the standard error of loss reserve. Taylor and Ashe (1983) introduces the second moment of estimates of outstanding claims. Their data is used in many other papers including this paper. Hayne (1985) provides an estimate of statistical variation in development factor methods when a lognormal distribution is assumed for those factors. Verrall (1991) derives unbiased estimates of total outstanding claims as well as an estimate of the variance of the estimate of expected total outstanding claim. England and Verrall (1999) provides analytic and bootstrap estimates of prediction errors in claims reserving. De Alba (2002) presents a Bayesian approach to obtain point estimates, probability intervals and other summary measures, such

as variance and quantiles. Han and Gau (2008) provide closed-form solutions for unbiased estimates of reserves and their corresponding standard errors by assuming lognormally distributed development factors.

Even with so many sophisticated methods developed over years, the Chain-Ladder method, a conceptually simple method for investigating the amount of loss reserve, is still the most commonly used methodology in the insurance industry. However, the Chain-Ladder method only provides a point estimate of the total reserve and does not provide any information about the variation inherent in the loss reserving process. That is, the Chain-Ladder method speaks in absolutes rather than probabilistic terms.

In order to speak in probabilistic terms, we propose a random selection method in the loss reserving process. By choosing appropriate weights, the expected total loss reserve suggested by the proposed random selection method is statistically equivalent to the expected total loss reserve suggested by the Chain-Ladder method. In addition, the proposed random selection method is able to use the Monte Carlo simulation technique for obtaining an estimate of standard error associated with the estimation of loss reserve.

Many currently existing loss reserve models can be categorized as parametric models. The choice of parametric model mostly depends on experience gained through analysis of data. The problem is that the result of a parametric model might be too rigid, and the rigidity of parametric approach can be overcome by removing the restriction that the model is from a parametric family. The motivation is to let the data speak for themselves. Nevertheless, nonparametric and parametric models are not mutually exclusive competitors. A nonparametric approach provides an alternative for actuaries. When there is no reasonable and adequate parametric model, a nonparametric model provides us the first look at the loss reserve estima-

tion.

The rest of this article is organized as follows. Section 2 introduces the loss reserving process with the Chain-Ladder method. Section 3 presents the loss reserving process with the proposed random selection method. Section 4 applies the kernel smoothing technique to obtain an improved confidence bands for the estimate of loss reserve. Section 5 applies the proposed method to a dataset used by Taylor and Ashe (1983) and a dataset used by Wisner (2001). Some remarks are made in section 6.

2 The Chain-Ladder Method

The Chain-Ladder method is the most commonly used methodology in determining loss reserves. In general, the loss reserve estimation begins with payments actually made for a certain line of business at successive development years.

Accident Year	Development Year								
	1	2	3	4	...	n-3	n-2	n-1	n
1	x_{11}	x_{12}	x_{13}	x_{14}	...	$x_{1,n-3}$	$x_{1,n-2}$	$x_{1,n-1}$	$x_{1,n}$
2	x_{21}	x_{22}	x_{23}	x_{24}	...	$x_{2,n-3}$	$x_{2,n-2}$	$x_{2,n-1}$	
3	x_{31}	x_{32}	x_{33}	x_{34}	...	$x_{3,n-3}$	$x_{3,n-2}$		
4	x_{41}	x_{42}	x_{43}	x_{44}	...	$x_{4,n-3}$			
⋮	⋮	⋮	⋮	⋮	...				
n-3	$x_{n-3,1}$	$x_{n-3,2}$	$x_{n-3,3}$	$x_{n-3,4}$...				
n-2	$x_{n-2,1}$	$x_{n-2,2}$	$x_{n-2,3}$...				
n-1	$x_{n-1,1}$	$x_{n-1,2}$...				
n	$x_{n,1}$...				

Table 1: **Incremental Loss Payments by Development Year**

Table 1 shows a typical data that the actuary will face in developing the loss reserve. It represents incremental loss payments by development years. Much of this data will be incomplete in the sense that the final claim cost will not be known, and a claim cost estimate must be used.

In addition, let us define i as the index for the accident year and j as the index for the development year, where $i = 1, \dots, n$ and $j = 1, \dots, n$. The entry for the accident year N and the development year 1 shows all dollars paid in the calendar year N on claims with an accident date in the year N . Similarly, the entry for the accident year N and the development year j shows the dollars paid in the calendar year $N + j - 1$ on claims with an accident date in the year N . Entries that have the same sum of i and j show the dollars paid in the same calendar year. This happens when entries are on the diagonal from the top left to the bottom right of Table 1, with the most recent accident date in the year n .

The second stage of the process is to transform Table 1 into cumulative payments through development years. This transformation is displayed in Table 2, in which

$$s_{ij} = \sum_{t=1}^j x_{it} \text{ for } i = 1, \dots, n; j = 1, \dots, n; \text{ and } i + j \leq n + 1. \quad (1)$$

Accident Year	Development Year								
	1	2	3	4	...	n-3	n-2	n-1	n
1	s_{11}	s_{12}	s_{13}	s_{14}	...	$s_{1,n-3}$	$s_{1,n-2}$	$s_{1,n-1}$	s_{1n}
2	s_{21}	s_{22}	s_{23}	s_{24}	...	$s_{2,n-3}$	$s_{2,n-2}$	$s_{2,n-1}$	
3	s_{31}	s_{32}	s_{33}	s_{34}	...	$s_{3,n-3}$	$s_{3,n-2}$		
4	s_{41}	s_{42}	s_{43}	s_{44}	...	$s_{4,n-3}$			
⋮	⋮	⋮	⋮	⋮	...				
n-3	$s_{n-3,1}$	$s_{n-3,2}$	$s_{n-3,3}$	$s_{n-3,4}$...				
n-2	$s_{n-2,1}$	$s_{n-2,2}$	$s_{n-2,3}$...				
n-1	$s_{n-1,1}$	$s_{n-1,2}$...				
n	s_{n1}				...				

Table 2: **Cumulative Loss Payments through Development Years**

From the cumulative payments, we calculate the age-to-age loss-development factors (sometimes called link ratios) as shown in Table 3. Each entry is the ratio of

successive development year cumulative payments, that is,

$$l_{ij} = \frac{s_{i,j+1}}{s_{ij}} \text{ for } i = 1, \dots, n-1; j = 1, \dots, n-1; \text{ and } i+j \leq n. \quad (2)$$

Accident Year	Ratio of Successive Development Years								
	1	2	3	4	...	n-3	n-2	n-1	n
1	l_{11}	l_{12}	l_{13}	l_{14}	...	$l_{1,n-3}$	$l_{1,n-2}$	$l_{1,n-1}$	
2	l_{21}	l_{22}	l_{23}	l_{24}	...	$l_{2,n-3}$	$l_{2,n-2}$		
3	l_{31}	l_{32}	l_{33}	l_{34}	...	$l_{3,n-3}$			
⋮	⋮	⋮	⋮	⋮	...				
n-3	$l_{n-3,1}$	$l_{n-3,2}$	$l_{n-3,3}$...				
n-2	$l_{n-2,1}$	$l_{n-2,2}$...				
n-1	$l_{n-1,1}$...				
n									

Table 3: Age-to-Age Loss-Development Factors Based on Cumulative Payments

When there is no further information available, we will assume that all loss-development factors beyond those given for the oldest accident year are equal to 1. That is, for the oldest accident year, there is no loss development beyond the last development year. After having the loss-development factors as seen in Table 3, the Chain-Ladder method develops the loss reserve via the following four-step process.

1. The single age-to-age column factors are chosen to model the loss development indicated by existing experience data.
2. The selected patterns of loss development are then projected to create the lower half of the loss-development triangle, so the model can be used to estimate the expected ultimate payment for each accident year.
3. The expected ultimate payment less the paid-to-date payment represent the reserve requirement for each accident year.

4. Finally, the total reserve (TR) is equal to the sum of reserve requirements for all accident years.

In order to find a single age-to-age loss-development factor for those unknown factors in each column of Table 3, one can use the arithmetic average, five-year average, or weighted average of known factors in each column. For the weighted average, the accident years with more losses and more recent data normally are assigned more weights. The well-known Chain-Ladder method is the result of having accumulated losses as weights. For illustrative purpose, we use this weighted average to create the lower half of the loss-development triangle. That is, we have the estimated age-to-age loss-development factor for the j^{th} column, l_j , defined as

$$l_j = \sum_{i=1}^{n-j} w_{ij} \cdot l_{ij}, \quad (3)$$

where

$$w_{ij} = \frac{s_{ij}}{\sum_{i=1}^{n-j} s_{ij}}, \quad j = 1, 2, \dots, n - 1;$$

and l_n is assumed to be 1.

After completing the lower half of loss-development triangle, the expected ultimate payment is calculated for each accident year. This is the expected cumulative payment at the end of development year n in Table 2. In these calculations, it is assumed that for each accident year all payments are made by the end of development year n . In order to calculate the expected ultimate payments, we need the amount of payments by the current date (losses paid-to-date) and the product of projected loss-development factors for each accident year (ultimate completion factors) as shown in Table 4.

Here, the paid-to-date payments are values on the diagonal of Table 2 and the

Accident Year	Paid-to-date	Completion Factor
1	s_{1n}	l_n
2	$s_{2,n-1}$	$l_{n-1} \times l_n$
\vdots	\vdots	\vdots
$n - 2$	$s_{n-2,3}$	$\prod_{j=3}^n l_j$
$n - 1$	$s_{n-1,2}$	$\prod_{j=2}^n l_j$
n	s_{n1}	$\prod_{j=1}^n l_j$

Table 4: **Paid-to-date and Completion Factors**

completion factor is the product of projected loss-development factors in each row of Table 3. With this information, the expected ultimate loss is calculated by the paid-to-date multiplies the cumulative development factor (also known as completion factor); and the estimated loss reserve for each accident year is equal to the expected ultimate loss minus the paid-to-date. Then, the total reserve for this block of business is determined by adding the estimated loss reserves for all accident years.

3 Random Selection

The Chain-Ladder method introduced in section 2 only provides a point estimate of the total reserve. It does not provide any information about the variation inherent in the process. What is the probability that the total reserve will exceed 10 million dollars? We are not able to answer this type of questions based on the results from this method. But, in practice, the Chain-Ladder method has been proven to be a valuable methodology. Therefore, we intend to develop a simulation technique that is as good as the Chain-Ladder method and at the same time allows us to speak in probabilistic terms rather than absolutes.

3.1 Selection Scheme

The basic assumption of the Chain-Ladder method is that the past is sufficiently indicative of the future. Thus, it motivates us to use the empirical distribution of loss-development factors from past accidental years to predict future loss-development factors. The first three stages of the process are exactly the same as the Chain-Ladder method. That means we will start with the payments actually made for a certain line of business at progressive development durations as shown in Table 1. Then, the data will be transformed into cumulative payments through development years as seen in Table 2. From Table 2, the age-to-age loss-development factors will be calculated the same way as we did for Table 3.

After having the loss-development factors, the proposed random selection method develops the loss reserve through the following iterative four-step process.

1. For each development year (i.e., by column) in Table 3, the unknown single age-to-age loss-development factors are randomly selected with replacement according to the weights assigned from a sample of known loss-development factors indicated by the upper half of the loss-development triangle.
2. The selected patterns of loss development are then projected to create the lower half of the loss-development triangle, so the model can be used to estimate the expected ultimate payment for each accident year.
3. The expected ultimate payment less the paid-to-date payment represent the reserve requirement for each accident year.
4. Finally, the total reserve (TR) is equal to the sum of reserve requirements for all accident years.

It can be shown that the total reserve obtained from the four-step random selection process is an unbiased estimator of the total reserve obtained from the Chain-Ladder method. To avoid confusion, we define L_j as a random variable that is randomly selected with replacement from the observed loss-development factors in the development year j . We also assume that L_1, L_2, \dots, L_{n-1} are independent of each other. The impact of this assumption will be further discussed in the later section.

Proposition 1 *The expected total loss reserve suggested by the proposed random selection scheme with weights given by (3) is statistically equivalent to the total loss reserve suggested by the Chain-Ladder method.*

Proof. The unknown single age-to-age loss-development factors are randomly selected from the observed loss-development factors. For the development year j (the j^{th} column) in Table 3, L_j is randomly selected, according to weights given by (3), with replacement from possible values of $\{l_{1j}, \dots, l_{n-j,j}\}$. The unknown entry l_{ij} , $n+1 \leq i+j \leq n+j$, in Table 3 is estimated by L_j . We have

$$\begin{aligned}
 E[L_j] &= \sum_{i=1}^{n-j} w_{ij} \cdot l_{ij} & (4) \\
 &= \sum_{i=1}^{n-j} \left[\frac{s_{ij}}{\sum_{i=1}^{n-j} s_{ij}} \cdot \frac{s_{i,j+1}}{s_{ij}} \right] \\
 &= \frac{\sum_{i=1}^{n-j} s_{i,j+1}}{\sum_{i=1}^{n-j} s_{ij}},
 \end{aligned}$$

for $j = 1, 2, \dots, n-1$. The equivalency can also be proved in a similar fashion for other cases by assigning probability accordingly. ■

The analytical variance can also be derived based on the proposed method. The ultimate completion factor for each year is defined as the product of the projected

loss-development factors for each accident year (i.e., each row) in Table 3. For example, for accident year n , the ultimate completion factor is defined as

$$L_{ult} = L_1 L_2 \cdots L_{n-1} L_n, \quad (5)$$

where L_{ult} is the ultimate completion factor and L_n is assumed to be 1. Since $L_1, L_2, \dots, \dots, L_{n-1}$ are independent, we have

$$\begin{aligned} Var(L_1 L_2 \cdots L_{n-1} L_n) &= E[(L_1 L_2 \cdots L_{n-1} L_n)^2] - [E(L_1 L_2 \cdots L_{n-1} L_n)]^2 \\ &= E[L_1^2] \cdots E[L_n^2] - [E(L_1) \cdots E(L_n)]^2. \end{aligned}$$

The first moments can be found in (4), and the second moments are given by

$$\begin{aligned} E[L_j^2] &= \sum_{i=1}^{n-j} w_{ij} \cdot l_{ij}^2 \\ &= \sum_{i=1}^{n-j} \left[\frac{s_{ij}}{\sum_{i=1}^{n-j} s_{ij}} \cdot \frac{s_{i,j+1}^2}{s_{ij}^2} \right] \\ &= \frac{\sum_{i=1}^{n-j} [s_{i,j+1}^2 / s_{ij}]}{\sum_{i=1}^{n-j} s_{ij}}. \end{aligned}$$

Furthermore, the variance of ultimate payment at the end of development year n can be found as

$$Var(s_{n1} \times L_1 L_2 \cdots L_{n-1} L_n) = s_{n1}^2 \times Var(L_1 L_2 \cdots L_{n-1} L_n), \quad (6)$$

where s_{n1} is the paid-to-date payment for the accident year n . Similarly, we can derive the variances of ultimate payments at the end of development year n for other accident years. But, there is no simple formula to determine the variance of total loss reserve due to the dependency among the ultimate payments.

Meanwhile, by sampling with replacement, the total number of arrangements for the lower triangle in Table 3 is $\prod_{k=1}^{n-1} k!$. For example, a 5 by 5 table in Table 3 will have 34,560 arrangements, and a 10 by 10 table in Table 3 will have 6.658607×10^{27} arrangements. It will be more efficient to implement the proposed method through the Monte Carlo simulation, in which we treat the original sample of values (the upper loss-development factor triangle in Table 3) as a stand-in population, reselect a value with replacement repeatedly, and recompute the reserves for each iteration. At the end of this random selection scheme, we will have thousands of simulated loss-development factor matrices. Thus, it enables us to estimate the distributions of desired quantities for a method that is based on the loss-development factors. One important application is to use this approach in the BF method. An example of the BF method will be demonstrated in section 5.2.

In summary, we use a weighted empirical distribution for each development year to randomly generate future loss-development factors in the lower right half of the loss-development matrix and compute one total reserve from each simulated matrix. Each simulation gives possible realizations of future loss-development factors, and the distribution of desired quantity can be estimated through thousands of such simulations.

3.2 Distribution of Total Reserve

In this section, we focus on the distribution of total reserve. For each iteration, we will complete the lower half of loss-development triangle. The expected ultimate payment is calculated for each accident year. Then, the total reserve for this block of business is determined by adding the estimated loss reserves for all accident years. When there are a large number of iterations, we will be able to approximate probabilistic

characteristics of the loss reserve for each accident year and the total loss reserve.

With each simulated loss-development matrix, we make an estimate of total reserve and label this estimate as \widehat{TR}_b . After we have taken B matrices, we approximate the standard error as

$$\sqrt{\frac{1}{B-1} \sum_{b=1}^B \left(\widehat{TR}_b - \overline{\widehat{TR}} \right)^2}, \quad (7)$$

where $\overline{\widehat{TR}}$ is the average of B total reserve estimates. The same idea can also be applied to determine the standard error of loss reserve for each accident year.

Overall, the random selection method is very flexible in terms of inputs. The actuary can decide the weight allocated to each observed development factor based on the available information and judgment. This method can be easily implemented in other loss reserving approaches where a similar development structure is used. For instance, the completion factors used in the health industry can be simulated similarly to obtain the distribution of total reserve. However, as promising as this method sounds, there are several important restrictions that need to be considered.

First of all, the Chain-Ladder method is considered to be somewhat biased in estimating the actual total reserve (Stanard, 1985). Using the random selection scheme with weights defined in (3), the resulting distribution of total reserve has a mean that is equivalent to the total reserve obtained from the Chain-Ladder method. Thus this proposed method does not provide a correction of the potential bias from the Chain-Ladder method. The standard error approximated by (7) provides an estimate of the variation of total reserve from the Chain-Ladder method, not necessarily the mean square error of the total reserve estimator. Nonetheless, having this type of information helps greatly in understanding the potential risk related to the business.

Secondly, the Chain-Ladder method has to be used with caution for early development years. The method could provide a very volatile result when the accumulated losses (paid or incurred) are relatively small comparing with the expected ultimate losses. In the extreme scenario where the accumulated loss for the first development year is zero, the loss-development factor becomes undefined, and the Chain-Ladder method can not be used. In general, if the early losses are not good approxy to the expected losses for the period, the Chain-Ladder method may provide a biased estimation (Halliwell, 2007). For many business lines, such as automobile insurance and homeowner insurance, losses are recognized quickly so that they become good approximations of expected total losses shortly after. Under these situations, the random selection scheme works best and is able to provide reliable results. For other business lines, such as medical malpractice and general liability, losses may not be recognized years after occurrence, and the early loss-development factors can be very volatile. An extra ounce of caution is needed under these circumstances, and alternative methods may be necessary for early development years.

Thirdly, the loss reserving methods through the loss-development triangle almost always face the fact that the data size is generally small. This is because there are many internal and environmental factors that can affect the loss reporting system over the years and the most recent years of data are more indicative of the future loss development. Even though the size of observed development factors for each development year is lower than the typical requirement for a reliable result from the bootstrap method (Chernick, 1999, page 150-151), the proposed method is different from the bootstrap method in several ways. The past experience is considered to be facts and will not be resampled, so the upper half of Table 3 remains unchanged through the entire process. Then, the future loss-development factors are randomly selected from the empirical distribution of observed sample. For the development year

j , we randomly select, with replacement, j future loss-development factors among $n - j$ observed data points. Each simulated matrix contains randomly selected loss-development factors in the lower-right triangle. If the number of policy years and development years are both n , the total number of randomly selected factors is $n(n - 1)/2$. For instance, in the case of $n = 10$, there are 9 data points available for the selection of 1 future loss-development factor in the 1st development year, 8 data points for the selection of 2 future loss-development factors in the 2nd development year, ..., and 1 data points for the selection of 9 future loss-development factors in the 9th development year. The total number of randomly selected future loss-development factors is 45. The chance of repeating the same lower-right triangle in the random selection method is very small.

Moreover, one of the key assumptions in the Chain-Ladder method is that the pattern of loss-development factors is indicative of the future loss development. Thus, it is assumed that the observed loss-development factors are somewhat representative and can be used to predict the future loss development. The actuary needs to validate the reasonability of this assumption and make adjustments as necessary before applying the method. When this assumption is seriously violated and cannot be fixed, alternative loss reserving methods are recommended.

In addition, there are relatively more data points at early development years. As the losses progress toward the ultimate losses, the size of data points decreases in algebraic order, while the loss-development factors are in general heading toward 1 in a much faster fashion. Therefore, for a short tailed or less volatile line of coverages, the lack of large sample size for later development years is mitigated by the stability of loss-development factors. Generally, for these business lines, most variation in the loss reserving comes from the early development years.

Lastly, the resulting distribution using the Monte Carlo simulation is discrete in

essence and may not be smooth, especially for the later development years. Smoothing techniques can be applied to solve this problem and obtain a better estimation. Additionally, the Monte Carlo simulation generates future loss-development factors based on the empirical distribution of observed loss-development factors for each development year. It is known that the variance of empirical distribution underestimates the actual variance by a factor of $\frac{n-1}{n}$, provided n is the number of observed loss-development factors. Most kernel smoothing techniques can mitigate this effect and give a more conservative variance estimation.

4 Kernel Smoothing

An inherent drawback of the Monte Carlo simulation in determining the loss reserves, particularly with those development years close to the development year n , lies in the discrete nature of the empirical distribution function. In addition, development factors can be very large in practice. The random selection method proposed in the previous section limits the selection of loss-development factors for each development year from the minimum to the maximum of the column data. By doing so, we might exclude possible large factors with positive probabilities and result in underestimating the total reserve.

4.1 Smoothed Random Selection

The objective with kernel smoothing is to create a density function that will in some way approximate the empirical distribution. For the j^{th} development year in Table 3, we have a sample of loss-development factors, $\{l_{1j}, l_{2j} \cdots, l_{n-j,j}\}$. At each point l_{ij} , a density function (kernel) corresponding to that point is created, and this density

function is denoted as

$$k_{ij}(l).$$

For each l_{ij} , $k_{ij}(l)$ satisfies the requirements of a probability density function. The kernel smoothed density function estimator is then a finite mixture (or weighted average) of these separated density functions. The weight applied to $k_{ij}(l)$ is the assigned probability. In the Chain-Ladder method, the weights are given by w_{ij} in (3). Therefore, the resulting kernel density estimator of the density function for the j^{th} development year in Table 3 is given by

$$\widehat{f}_j(l) = \sum_{i=1}^{n-j} w_{ij} \cdot k_{ij}(l). \quad (8)$$

Thus, to obtain an improved estimate, we modified our simulation procedure as follows.

1. For j^{th} development year in Table 3, the unknown single age-to-age loss-development factors (i.e., random variable L_j) are randomly selected from the kernel smoothing estimator of the density function $\widehat{f}_j(l)$, where $j = 1, \dots, n-1$. And, the loss-development factor for development year n is assumed to be 1.
2. The selected patterns of loss development are then projected to create the lower half of the loss-development triangle, so the model can be used to estimate the expected ultimate payment for each accident year.
3. The expected ultimate payment less the paid-to-date payment represent the reserve requirement for each accident year.
4. Finally, the total reserve (TR) is equal to the sum of reserve requirements for all accident years.

Notice that step 2 to step 4 are the same steps as described in section 3.1. The advantage of using the kernel estimation in the first step is its flexibility. And, the first step can be achieved by the following two substeps.

- 1a. Draw a value at random from the weighted empirical distribution of each column in Table 3.
- 1b. Draw a value at random from the kernel whose mean is equal to the value drawn at the step 1a.

However, there are two important issues that need to be addressed. The first issue is how to select the kernel, and the second is the impact of bandwidth selection. These issues will be addressed in the following sections.

4.2 Selection of Kernel

The choice of kernel and the selection of bandwidth are closely related. The kernel $k(\cdot)$ is only well defined up to a scale. For example, the choices of $N(0, 1)$ and $N(0, 2)$ are identical choices. The optimal kernel in density estimation was discussed in Hodges and Lehmann (1956) and Epanechnikov (1969). In practice, the Epanechnikov kernel is referred as the optimal kernel; nevertheless, the choice of kernel is not critical. This conclusion is based on the efficiencies of several kernels compared to the optimal kernel shown in Table 5.

The key message is that the suboptimal kernels lose very little in performance. And, these results suggest that most unimodal kernel densities perform about the same as each other. There are three kernels discussed in Klugman et al. (2004), namely uniform kernel, triangular kernel, and Gamma kernel. We will discuss the implementation of these three kernels in S-PLUS.

Kernel	Efficiency
Epanechnikov	1
Biweight	0.994
Triweight	0.987
Triangular	0.986
Normal	0.951
Uniform	0.930

Table 5: **Efficiencies of several kernels compared to the optimal kernel. (Source: Wand and Jones, 1995)**

4.2.1 Uniform Kernel

Let y be the randomly selected data point. The uniform kernel is given by

$$\begin{aligned}
 k_y(x) &= \begin{cases} 0, & x < y - b, \\ \frac{1}{2b}, & y - b \leq x \leq y + b, \\ 0, & x > y + b \end{cases} \quad (9)
 \end{aligned}$$

where b is the bandwidth parameter. The simulated observation is a random selection from a uniform distribution between $y - b$ and $y + b$. The determination of bandwidth will be discussed in the bandwidth selection section.

4.2.2 Triangular Kernel

Let y be the randomly selected data point. The triangular kernel is given by

$$\begin{aligned}
 k_y(x) &= \begin{cases} \frac{x-(y-b)}{b^2}, & y - b \leq x \leq y \\ \frac{(y+b)-x}{b^2}, & y \leq x \leq y + b, \\ 0, & \text{Otherwise} \end{cases} \quad (10)
 \end{aligned}$$

where b is the bandwidth parameters. The distribution function of the triangular kernel can be expressed as

$$K_y(x) = \begin{cases} 0, & x < y - b \\ \frac{(x-(y-b))^2}{2b^2}, & y - b \leq x \leq y \\ 1 - \frac{((y+b)-x)^2}{2b^2}, & y \leq x \leq y + b \\ 1, & x > y + b \end{cases}. \quad (11)$$

Therefore, one can simulate an observation from the triangular kernel by using the inverse transformation of the distribution function. This can be achieved by the first 9 lines of S-PLUS codes in the appendix A, in which, the input values of y and b are the selected loss-development factor and the pre-determined bandwidth, respectively.

4.2.3 Gamma Kernel

As mentioned previously, it is possible to have large loss-development factors with positive probabilities in practice. A Gamma kernel allows us to realize this possibility, since it is a density function with a long tail. Meanwhile, the Gamma kernel does not require choosing a bandwidth b . In general, if y is the randomly selected data point, the Gamma kernel is given by

$$k_y(x) = \frac{1}{\theta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\theta}}, x > 0, \quad (12)$$

where α is the shape parameter and

$$\theta = \frac{y}{\alpha}.$$

Notice that the expected value is equal to $\alpha\theta$, and the variance is equal to $\alpha\theta^2$. That is θ can be determined by the ratio of the variance and the mean. Thus, with the selected value of y and the pre-determined variance, the simulated observation is

performed with the built-in function `rgamma()` in S-PLUS.

4.3 Bandwidth Selection

The practical implementation of the kernel density estimator requires the selection of the bandwidth b . How best to choose bandwidth parameters is still an ongoing research. In this article, we intend to have a class of simple and easily computable formulas which will find a bandwidth that is reasonable. If we assume for simplicity the unknown density to be $N(\mu, \sigma)$. Then, a quick and simple bandwidth is suggested by the following rule of thumb (Härdle, 1991).

$$\begin{aligned}\hat{b} &= \left(\frac{4\hat{\sigma}^5}{3n}\right)^{1/5} \\ &\approx 1.06\hat{\sigma}n^{-1/5}.\end{aligned}\tag{13}$$

A more robust estimate is to use the interquartile range \hat{R} , which is defined as

$$\hat{R} = \text{the sample } 75^{\text{th}} \text{ percentile} - \text{the sample } 25^{\text{th}} \text{ percentile}.$$

The rule of thumb is then modified as

$$\hat{b} = 0.79\hat{R}n^{-1/5}.\tag{14}$$

A better rule of thumb is the combination of both rules above. That is,

$$\hat{b} = 1.06 \min\left(\hat{\sigma}, \frac{\hat{R}}{1.34}\right)n^{-1/5},\tag{15}$$

which can be computed by using `bandwidth.nrd()` S-PLUS function.

5 Applications

In this section, we present two examples to illustrate how the random selection method can be applied to find loss reserves and the corresponding standard errors. The first example compares results of the proposed method with results of others, and the data is taken from Taylor and Ashe (1983). The second example applies the proposed method to the BF method (Bornhuetter and Ferguson, 1972), and this example uses the data from Wiser (2001, page 257, Chapter 5, Foundations of Casualty Actuarial Science).

5.1 Random Selection: A Close Look

The data for this example has been used by many researchers. The accumulated incurred losses are shown in Table 6, and the age-to-age loss-development factors are displayed in Table 7. Since no further information is available from the data, we assume that the development to ultimate factor for the last observed development year is 1 and the accumulated claims are the paid-to-date values.

The reserve estimates are exhibited in Table 8, and the standard errors as a

358	1,125	1,735	2,218	2,746	3,320	3,466	3,606	3,834	3,901
352	1,236	2,170	3,363	3,809	4,130	4,658	4,924	5,349	
291	1,292	2,219	3,235	3,986	4,133	4,629	4,909		
311	1,419	2,195	3,757	4,030	4,382	4,588			
443	1,136	2,128	2,898	3,403	3,873				
396	1,333	2,181	2,986	3,692					
450	1,297	2,429	3,492						
359	1,421	2,864							
377	1,363								
344									

Table 6: **Run-Off Triangle (Accumulated incurred loss, by 1,000)**

3.1432	1.5428	1.2783	1.2377	1.2092	1.0441	1.0404	1.0630	1.0177
3.5106	1.7553	1.5500	1.1325	1.0843	1.1278	1.0571	1.0863	
4.4485	1.7167	1.4583	1.2321	1.0369	1.1200	1.0606		
4.5680	1.5471	1.7118	1.0725	1.0874	1.0471			
2.5642	1.8730	1.3615	1.1742	1.1383				
3.3656	1.6357	1.3692	1.2364					
2.8843	1.8720	1.4378						
3.9533	2.0157							
3.6192								

Table 7: **Age-to-Age Development Factors**

percentage of estimated reserves are given in Table 9. The random selection Chain-Ladder method is based on (3). The results of random selection with equal weights (i.e., $w_{ij} = \frac{1}{n-j}$) is also provided here.

Policy	Random Selection	Random Selection	Uniform	Triangular	Gamma	Chain-Ladder	Verrall	Taylor
Year	Chain-Ladder using (3)	(with equal weight)	Kernel	Kernel	Kernel	Mack (1993)	(1991)	Ashe
2	95	95	95	95	94	95	96	298
3	469	459	460	459	460	470	439	600
4	709	694	693	694	692	710	608	745
5	986	965	963	964	960	985	1,011	1,077
6	1,411	1,430	1,436	1,430	1,433	1,419	1,423	1,788
7	2,182	2,226	2,231	2,243	2,243	2,178	2,150	2,879
8	3,917	3,966	3,957	3,959	3,933	3,920	3,529	4,221
9	4,277	4,309	4,293	4,291	4,277	4,279	4,056	4,866
10	4,608	4,738	4,746	4,738	4,726	4,626	4,340	5,827
Overall	18,865	18,881	18,874	18,873	18,818	18,681	17,652	22,301

Table 8: **Estimated Reserves, In 1000s**

In all cases, 10,000 loss-development factor matrices are simulated. The results for the “Chain-Ladder Mack”, “Taylor and Ashe”, and “Verrall (1991)” are taken from Mack (1993). In practice, the Chain-Ladder method is not able to provide an estimation of standard error. The standard errors for the Chain-Ladder method shown in Table 9 are approximated by Mack (1993).

The reserve estimates of random selection methods are literally the same as re-

Policy	Random Selection	Random Selection	Uniform	Triangular	Gamma	Chain Ladder	Verrall	Taylor
Year	Chain-Ladder using (3)	(with equal weight)	Kernel	Kernel	Kernel	Mack (1993)	(1991)	Ashe
2	0.0%	0.0%	26.1%	9.2%	9.3%	80.0%	49.0%	27.0%
3	12.3%	13.7%	14.7%	13.3%	28.9%	26.0%	37.0%	20.0%
4	10.0%	10.4%	11.7%	10.9%	21.4%	19.0%	30.0%	18.0%
5	18.9%	19.3%	22.1%	20.7%	31.6%	27.0%	27.0%	16.0%
6	22.9%	23.2%	25.4%	24.3%	36.0%	29.0%	25.0%	16.0%
7	21.7%	21.6%	23.5%	22.3%	32.4%	26.0%	25.0%	14.0%
8	21.2%	21.3%	23.2%	22.1%	32.4%	22.0%	27.0%	14.0%
9	20.1%	20.4%	21.7%	21.0%	30.2%	23.0%	30.0%	14.0%
10	25.8%	25.3%	27.3%	26.1%	38.0%	29.0%	38.0%	14.0%
Overall	9.6%	9.6%	10.4%	10.0%	14.5%	13.0%	15.0%	9.0%

Table 9: **Standard Error in Percentage of Estimated Reserves**

sults from the traditional Chain-Ladder model. And, the proposed random selection method does perform as we anticipated. It provides us the total reserve that is statistically equivalent to the total reserve suggested by the traditional Chain-Ladder method and the variation inherent in the process.

For simplicity, the rest of discussion and comparison will be based on random selection methods with equal weights. The uniform and triangular kernel smoothing estimators are based on smoothed random selection methods introduced in section 4 and are based on the bandwidth selector given by (15). A complete S-PLUS program for the triangular kernel with equal weights is provided in the appendix A.

In addition, the result of using the Gamma kernel is also shown in this example. Unlike the other two kernels, the Gamma kernel requires a pre-determined variance. We use the sample variance of all column data in Table 7. This is a very conservative estimation by assuming each kernel has the same variance as the entire empirical distribution. Since the Gamma kernel allows the possibility of large loss-development factors during the selection, a large variation is expected from this process. The estimated distribution of total reserve is pictured in Figure 1.

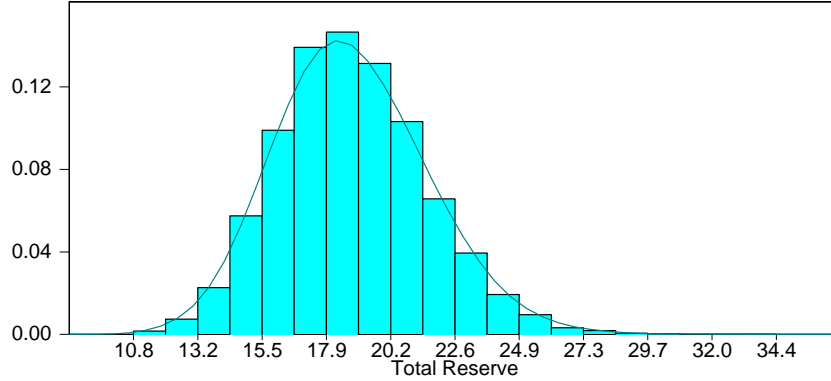


Figure 1: **Estimated Total Reserve Distribution with Gamma Kernel (in millions)**

Because the random selection process provides the distribution of total reserve through simulations, many interested statistics can be quickly obtained. For instance, the principal-based reserve may require a conditional tail expectation (CTE) at a certain level that is set by regulators. The reader is referred to Brazauskas et al. (2008) for further discussions of CTE. The 95% CTE of total reserve can be approximated by taking the average of the highest 5% of simulated total reserves. Table 10 lists the calculated 95% CTEs under different kernel smoothing functions.

Kernel	95% CTE
None	22,907,178
Uniform	23,204,998
Triangular	23,138,165
Gamma	24,931,412

Table 10: **95 percent CTE**

Meanwhile, based on the triangular kernel, a sensitivity test on the selection of bandwidth is conducted to show its impact on the total reserve. The results

Bandwidth	Total Reserve	Standard Error	Standard Error/ Total Reserve	95 Percentile Confidence Interval
b	18,873	1,878	10%	(15,412, 22,845)
2b	18,870	2,082	11%	(15,037, 23,197)
3b	18,880	2,397	13%	(14,492, 23,846)

Table 11: **Sensitivity Test on Bandwidth with Triangular Kernel In 1,000s**

are presented in Table 11. The bandwidth b in Table 11 is based on (15). As we expected, the estimation of total reserve remains practically the same in all cases. The standard error increases when the bandwidth becomes wider. A 95th percentile confidence interval is also provided in Table 11.

In addition, during the simulation process, we have implicitly assumed that loss-development factors for the same accident year are independent to each other. However, there might be some correlation between loss-development factors in adjacent development years. To examine the effect of independence assumption, the following example is studied under a bivariate normal structure with a correlation coefficient factor.

Suppose the loss-development factor random variable L_j follows a log-normal distribution with parameters μ_j and σ_j^2 . That is $\log(L_j) \sim Normal(\mu_j, \sigma_j^2)$, and the estimated μ_j and σ_j are given in Table 12 (see Han and Gau, 2008).

j	1	2	3	4	5	6	7	8	9
μ_j	1.2557	0.5530	0.3694	0.1649	0.1041	0.0807	0.0513	0.0720	0.0176
σ_j	0.1785	0.0903	0.0888	0.0535	0.0521	0.0362	0.0084	0.0108	0.0084

Table 12: **Estimated Parameters**

A bivariate normal structure on the log-development factors can be implemented to check the effect of the removal of the independence assumption. In details, we assume the joint distribution of $\log(L_j)$ and $\log(L_{j+1})$ is a bivariate normal with cor-

relation coefficient ρ . Thus, the conditional distribution of $\log(L_{j+1})|\log(L_j)$ follows a normal distribution. That is,

$$\log(L_{j+1})|\log(L_j) \sim Normal(\bar{\mu}_j, \bar{\sigma}_j^2), \quad (16)$$

where $\bar{\mu}_j = \mu_{j+1} + \rho \frac{\sigma_j}{\sigma_{j+1}}(\log(L_j) - \mu_j)$ and $\bar{\sigma}_j^2 = \sigma_{j+1}^2(1 - \rho^2)$. For each fixed ρ , the random selection process (each with 1000 simulations) is repeated 100 times to obtain 100 estimates of the mean and standard error of the total reserves. The average of these 100 estimates are summarized in Table 13.

Correlation Coefficient ρ	-0.5	-0.25	0	0.25	0.5
Average Mean Total Reserve	18,845	19,059	18,733	18,921	19,115
Average Standard Error	1,562	1,580	1,592	1,610	1,610

Table 13: **The Impact of Correlation Structure**

The addition of correlation structure causes some small variation in the average of 100 mean total reserve estimates. The average of 100 standard error estimates increases slightly as the correlation coefficient increases. Overall, the effect of independence assumption is insignificant in this example. This effect could be more significant if the observed development factors for the most recent accident years lie in the tail side of the distribution. For instance, a large development factor of the most recent accident year together with a large positive ρ could cause the total loss reserve significantly higher.

5.2 Random Selection and BF Method

The application of random selection method to the BF method is demonstrated in this example. The loss development data, including expected ultimate losses, is given in Table 14, and the age-to-age loss-development factors are shown on the run-off

Accident Year	Ages in Months							Initial Expected Loss
	12	24	36	48	60	72	84	
1994	58,641	74,804	77,323	77,890	80,728	82,280	82,372	81,557
1995	63,732	79,512	83,680	85,366	88,152	87,413		89,654
1996	51,779	68,175	69,802	69,694	70,041			78,237
1997	40,143	67,978	75,144	77,947				79,505
1998	55,665	80,296	87,961					83,738
1999	43,401	57,547						65,934
2000	28,800							45,004

Table 14: **Cumulated Incurred Losses**

Accident Year	Ages in Months							
	12-24	24-36	36-48	48-60	60-72	72-84	84-Ult	
1994	1.276	1.034	1.007	1.036	1.019	1.001	1.01	
1995	1.248	1.052	1.020	1.033	0.992		1.01	
1996	1.317	1.024	0.998	1.005			1.01	
1997	1.693	1.105	1.037				1.01	
1998	1.442	1.095					1.01	
1999	1.326						1.01	
2000							1.01	

Table 15: **Age-to-age Loss Development Factors**

triangle in Table 15. The loss-development factor from 84 months to ultimate is assumed to be 1.01 for all accident years.

The Bornhuetter–Ferguson (BF) method (Bornhuetter and Ferguson, 1972) uses the combination of development factors and expected losses to estimate reserves. The main advantage of the BF method is that it reduces the impact of unexpected losses and therefore increases the stability of the reserve estimate. The proposed method can be applied to estimate the standard errors of BF reserve estimates. For accident year $i = 1994, \dots, 2000$, suppose the expected loss from other resources is EL_i . Then the estimated reserve for policy year i would be $EL_i(1 - 1/L_i^{ult})$, where L_i^{ult} is the cumulative development factor for i^{th} accident year.

The random selection method is applied for 10,000 simulations. Table 16 records

Accident Year	Random Selection Chain-Ladder		Random Selection BF		BF (Wiser, 2001) Estimated Reserve
	Mean	Stdev	Mean	Stdev	
1994	4,972	0	4,955	0	4,964
1995	7,088	0	7,102	0	7,023
1996	4,756	955	4,853	1,032	4,421
1997	18,856	1,535	18,761	1,450	17,906
1998	30,311	2,191	29,744	1,865	29,232
1999	31,157	2,472	31,205	2,235	32,616
2000	32,847	4,694	32,832	2,817	33,475
Total	129,987	5,980	129,454	4,497	129,637

Table 16: **Estimated Reserves**

the statistics of the simulated distribution of total reserve for each accident year. It confirms that the estimate of total reserve for each accident year from the random selection BF method is very similar to the results obtained from the random selection Chain-Ladder method and the traditional BF method. The random selection method provides an extra information on the variation of each estimate. Moreover, the random selection BF method produces a more stable result as expected.

6 Remarks

The proposed random selection method is a nonparametric method that allows the data to speak for itself. For any given method that is in the family of loss-development triangle method, the random selection method produces an estimated distribution of total reserve. By assigning proper weights, this method can be applied to estimate the distribution of total reserve for the Chain-Ladder method and the BF method. This is a great enhancement over these two methods since they only produce point estimates of the total reserve.

The drawback of the random selection method introduced in section 3 lies in the discontinuous nature of empirical distribution function. Therefore, it is desirable

to generate the value of loss-development factor that should be used to model the tail of loss payment pattern from a smoothed kernel density. Meanwhile, as seen in examples, the most recent accident years as of the valuation date are critical since those accident years account for the majority of total claim reserves. For a more conservative reserve estimation, a kernel density with a longer (or flatter) tail should be considered. Furthermore, the actuary should recognize issues related to the choice of kernel and the selection of bandwidth. In practice, we recommend that the actuary uses a number of combinations between the choice of kernel and the selection of bandwidth to obtain a range of reasonable estimates. The actuary then uses his or her experience and judgement to provide an estimate of the loss reserve.

As emphasized in many literature, it is very important to understand the ability of a model and its limitations. Actuaries often apply multiple methods to make a more accurate estimation. Practitioners are welcomed to use the proposed random selection methodology as a sound reference in the reserving process and combine other knowledge to assess the reserves.

Acknowledgement

The authors gratefully acknowledge the referees for their extensive comments and suggestions that led to a significant improvement of the paper.

References

1. Bornhuetter, R.L. and Ferguson, R.E., 1972. The Actuary and IBNR. Proceedings of the Casualty Actuarial Society, 59:181-195.
2. Brazauskas, V, Jones, B.L., Puri, M.L., and Zitikis, R., 2008. Estimating Conditional Tail Expectation with Actuarial Applications in View. Journal of Statistical Planning and Inference, Volume 138, Issue 11, Pages 3590-3604

3. Brown, R.L. and Gottlieb, L.R., 2001. Introduction to Ratemaking and Loss Reserving for Property and Casualty Insurance. Second Edition. ACTEX.
4. Casualty Actuarial Society, 2001. Foundations of Casualty Actuarial Science. Fourth Edition.
5. Chernick, M.R., 1999. Bootstrap Methods: A Practitioner's Guide. John Wiley & Sons.
6. DE Alba, E., 2002. Bayesian Estimation of Outstanding Claim Reserves. North American Actuarial Journal, 6(4):1-20.
7. England, P. and Verrall, R., 1999. Analytic and Bootstrap Estimates of Prediction Errors in Claims Reserving. Insurance: Mathematics and Economics, 25:281-293.
8. Epanechnikov, V.A., 1969. Non-parametric Estimation of A Multivariate Probability Density. Theory Probab. Appl. 14:153-158.
9. Finger, R.J., 1976. Modelling Loss Reserve Developments. Proceedings of the Casualty Actuarial Society, 63:90-105.
10. Halliwell, L.J., 2007. Chain-Ladder Bias: Its Reason and Meaning. Variance, Volume 1/Issue 2: 214-247.
11. Han, Z. and Gau, W.-C., 2008. Estimation of Loss Reserves with Lognormal Development Factors. Insurance: Mathematics and Economics, 42: 389-395.
12. Härdle, W., 1991. Smoothing Techniques With Implementation in S. Springer-Verlag New York Inc.

13. Hayne, R.M., 1985. An Estimate of Statistical Variation in Development Factor Methods. *Proceedings of the Casualty Actuarial Society*, 72:25-43.
14. Hodges, J.L. and Lehmann, E.L., 1956. The Efficiency of Some Nonparametric Competitors to the t-test. *Ann. Math. Statist.* 13:324-335.
15. Klugman, S.A., Panjer, H.H., and Willmot, G.E., 2004. *Loss Models: From Data to Decisions, 2nd Edition*. Wiley.
16. Mack, T., 1993. Distribution-Free Calculation of The Standard Error of Chain Ladder Reserve Estimates. *ASTIN Bulletin*, 23(2): 213-225.
17. Stanard, J.N., 1985. A Simulation Test of Prediction Errors of Loss Reserve Estimation Techniques. *Proceedings of the Casualty Actuarial Society*, 72: 124-153.
18. Taylor, G.C. and Ashe, F.R., 1983. Second Moments of Estimates of Outstanding Claims. *Journal of Economics*, 23:37-61.
19. Taylor, G.C., 2000. *Claim Reserving. An Actuarial Perspective*. New York: Elsevier Science Publishers.
20. Verrall, R.J., 1991. On the Estimation of Reserves from Nonlinear Models. *Insurance: Mathematics and Economics*, 10(1):75-80.
21. Wand, M.P. and Jones, M.C., 1995. *Kernel Smoothing*. New York: Chapman and Hall.
22. Wiser, R.F., 2001. *Foundations of Casualty Actuarial Science*. *Casualty Actuarial Society* (Chapter 5).

Appendix A: S-PLUS Codes with Triangular Kernel

```
triangle <-  
function(y, b)  
{  
  u <- runif(1, 0, 1)  
  x <- sqrt(2 * u * b^2) + (y - b)  
  if(u >= 0.5)  
    x <- (y + b) - sqrt(2 * (1 - u) * b^2)  
  x  
}  
  
up.to.date.values  
<- c(3901463,5348785,4909315,4588268,3873311,  
      3691712,3492130,2864498,1363294,344014)  
development <- as.matrix(development)  
  
### Bandwidth Selection Using (17)  
for (k in 1:9)  
{  
  b[k]<- bandwidth.nrd(development[1:(10-k),k])/4  
}  
  
b[9]<-0.5*b[8]  
rs.tri<-numeric(9)  
rs.tri.development <- matrix(0,10,9)
```

```

product.development.factors.tri <- numeric(10)
product.development.factors.tri[1]<-1
reserves.tri <-matrix(0,10000,10)
total.reserve.tri <-numeric(10000)

for (m in 1:10000)
{

for (i in 1:9)
{
for (k in 1:9)
{
y<- development[ceiling(runif(1,0,length(development[,1])-k)),k]
bb<- b[k]
rs.tri[k]<-triangle(y, bb)
}
### Fill the random sample to the lower triangle ###
rs.tri.development[11-i,i:9]<-rs.tri[i:9]
### Fill in the observed loss-development factors to the upper triangle ###
rs.tri.development[i,1:(10-i)]<-development[i,1:(10-i)]
}

for (j in 2:10)
{
### product of development factors for each accident year ###
product.development.factors.tri[j]<-prod(rs.tri.development[j,9:(11-j)])
}
}

```

```
}
```

```
predicted.accumulated.claims.tri
```

```
<- up.to.date.values*product.development.factors.tri
```

```
reserves.tri[m,]<-predicted.accumulated.claims.tri-up.to.date.values
```

```
total.reserve.tri[m] <- sum(reserves.tri[m,])
```

```
}
```