# Anatomy of Actuarial Methods of Loss Reserving

Prakash Narayan, Ph.D., ACAS

**Abstract:** This paper evaluates the foundation of loss reserving methods currently used by actuaries in property casualty insurance. The chain-ladder method, also known as the weighted loss development method in North America, is the most commonly used actuarial technique for loss reserving and setting liabilities for property/casualty insurers. Many actuaries believe that the basic assumption underlying this model is the future development of losses is dependent on losses to date for each accident year. We shall see that this is not the case and the method may be rooted in the complete independence of future loss development. The alternative assumptions are, in this author's opinion, a more natural way of analyzing the loss triangle. We shall also show that most of the methods used by actuaries are based on one common basic model, and the differences lie in how and which of the parameters are being estimated. The exposition provides some new insight to reserving methods. While it enriches our understanding of the loss reserving process and defines the common thread among various methods, it challenges some commonly held views in the actuarial profession. The exposition here points out a flaw in the Bornhuetter-Ferguson methodology as well as questions the basic framework of the loss development methodology. We shall show that we can obtain the same results as the loss development method under the assumption that the future losses are independent of what we know currently.

We introduce a new method, termed the exposure development method, which has some advantages over traditional loss development methods in some situations. The proposed methodology allows us to construct several new estimators. One can estimate the ultimate losses by combining the information gleaned from paid losses and the incurred loss triangles. Most importantly, this methodology provides better analytical tools to examine the model, look for outliers, and provides an alternative method of estimating the variability of reserves.

## INTRODUCTION

The results presented in this paper are quite basic and there is no need to review the current state of knowledge to proceed. For brevity, it will be appropriate to refer to them as needed in our exposition. Let $X_{i,j}$ denote the losses paid for the accident year $i$ in the $j^{th}$ year of development, where $i, j = 1, 2 \ldots n$. We assume that we have observed $X_{i,j}$ for $i + j < n + 2$ and are interested in estimating $X_{i,j}$ for $i + j = n + 2, n + 3 \ldots 2n$. Once we have estimated these, we could add them and compute the ultimate losses. In this paper, we restrict our attention to the development period n and assume that the losses are fully developed by that time. Any development beyond period n is outside the scope of the results presented here. Although we will mainly focus on the paid loss triangle, the methodology presented here can equally be applied to incurred or reported loss triangles. We also assume that we have some information available about the exposure for each accident year. For example, the earned premium for each accident year may be known. Although any measure of exposure will suffice for our purpose. If we have prior information about the ultimate losses, that may be used as an exposure base as well and might possibly be the best exposure base. The ultimate losses are exposure times a rate, and they are identical if the loss rate is constant. Sometimes we have used these interchangeably and the author assumes that does not cause any misunderstanding. As we shall see, the assumed knowledge of exposures is for exposition of the ideas presented here and is not necessary. Let us denote $E_i$ be the exposure amount for the accident year *i*. We shall use the Buhlman (1967) method to estimate the average loss by development period.

We compute

$$r_j = \frac{\sum_{i=1}^{n-j+1} X_{i,j}}{\sum_{i=1}^{n-j+1} E_i}, j = 1, 2 \ldots n. \tag{1.1}$$

However, we do not need to compute $r_1$, so the number of parameters we need and use is only $n$ - 1. If we use earned premium as a proxy for the exposure, the method is known as the partial loss ratio method. One should note that this method does not assume any relationship between development periods. We estimate

$$\hat{X}_{i,j} = E_i \times r_j \text{ for } i + j > n + 1. \tag{1.2}$$

This method, although somewhat popular in Europe, is seldom used in North America. However, we shall see that this method can be used as the building block of the loss development method.

Now let us assume that the exposures $E_i$s are not known and we want to estimate them from the data itself. It will suffice for our purpose if we have the estimates of relative exposure levels for each accident year, and that information is sufficient to compute $r_j$ and hence the values of the unpaid losses, which is our primary goal. We assume that the exposure level for the first accident year is unity ($E_1$ = 1) and try to estimate the future accident years' exposure relative to the first accident year's exposure. We compute what we call exposure development factors (EDFs).

$$d_k = \frac{\sum_{i=1}^{k+1} \sum_{j=1}^{n-k} X_{i,j}}{\sum_{i=1}^{k} \sum_{j=1}^{n-k} X_{i,j}}. \tag{1.3}$$

It may be easy to relate these factors to weighted loss development factors. All we have done is changed the process of loss development from operating in columns to operating in rows.

Let us define

$$D_k = d_1 \times d_2 \times \ldots \times d_k. \tag{1.4}$$

$D_k$ is the estimated total earned exposure by accident year $k$ +1 relative to accident year 1.

These exposure development factors can then be used to estimate the relative individual accident year exposures. The exposure for accident year $k$ +1 relative to the first accident year is $D_k - D_{k-1}$.

We could use these estimated relative exposures to compute $r_k$ and then using equation (1.2) compute the unknown elements of the loss rectangle.

One should note that we have estimated 2(*n* - 1) parameters in the process, (*n* - 1) parameters for the exposure level and another (*n* - 1) parameters for the development period rates.

It is interesting to note that one need not compute the payment year rates. One can directly estimate the unobserved element by computing

$$\hat{X}_{i,j} = d_{i-1} \times \sum_{j=1}^{i-1} X_{i,j} - \sum_{j=1}^{i-1} X_{i,j}, \quad i + j > n + 1. \tag{1.5}$$

One can easily verify that the results so obtained are the same that one would obtain by the more elaborate procedure stated earlier. Similar to the loss development method, this requires computing only (*n* - 1) parameters. We will call this method the exposure development method. The exposure development method has its advantages over the loss development method and may be a better way of analyzing loss triangles, as we shall see further on. We have defined our computational scheme based on incremental loss data. For computational purpose, it may be better to use cumulative loss triangles as we do in the loss development method. The computational procedure for the exposure development method is similar to the weighted loss development method. The difference is that we first transpose the incremental loss triangle and use this triangle to compute the cumulative loss triangle and carry out the same computation as for the weighted loss development method.

A quite surprising observation is that the estimates so obtained are those that one would obtain if the weighted loss development method had been used. The proof is trivial and one can easily verify that the formula for estimating $X_{i,j}$ for the exposure development method is equivalent to the weighted loss development method, where the unobserved $X_{i,j}$ are estimated by the formula

$$\hat{X}_{i,j} = \sum_{k=1}^{j-1} X_{i,k} \times \frac{\sum_{l=1}^{i-1} \sum_{k=1}^{j} X_{l,k}}{\sum_{l=1}^{i-1} \sum_{k=1}^{j-1} X_{l,k}} - \sum_{k=1}^{j-1} X_{i,k}. \tag{1.6}$$

Where unobserved values of $X_{i,j}$ used in equation (1.6) are estimated first and then are treated as the observed values in the equation. The pictorial view shown in Figure 1 helps illustrate the approach better. The symbols A, B, C and D represent the sum of incremental losses of the area they cover. The right top formula in the figure 1, represents the estimate when weighted loss development method is used. The bottom left is the formula for exposure development, and the bottom right is the formula when we first estimated the exposure levels and then use Buhlman's method. We do not show the calculation of exposures (F in the formula in Figure 1) as it cancels out.

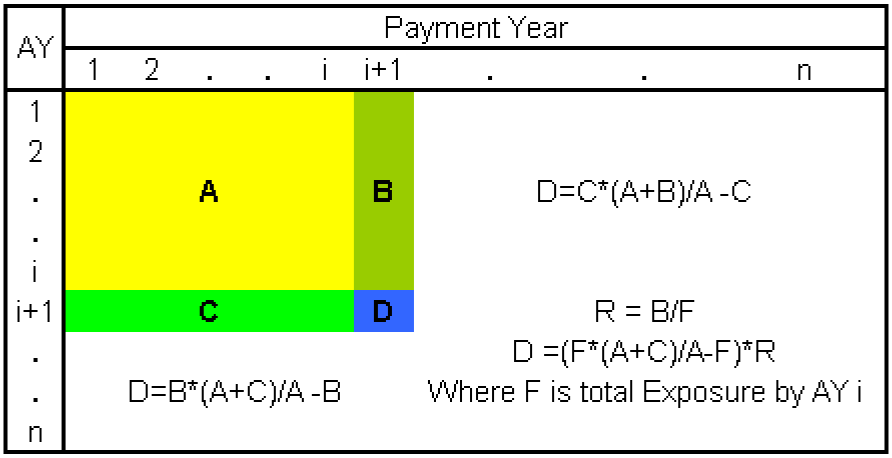

Figure1

The important point to note is that by using the alternate derivation (i.e., if we compute the relative exposures first and then use equation (1.2)) we have estimated 2(*n* - 1) parameters and arrive at the same answer as the weighted loss development method or the exposure development method, which appear to have (*n* - 1) parameters. The contrast in the number of parameters is puzzling. The only explanation I have come up with is based on our misunderstanding of what we are trying to estimate. The general belief that our aim in loss reserving is to find a number for the value of ultimate losses that will be paid when all the claims arising from that accident year are finally settled does not follow statistical logic. In a statistical framework, the ultimate losses are a random variable. A random variable cannot be estimated. The statistical methods are not meant to estimate a random outcome or the results of a flip of a coin. All one can do is to estimate the parameters associated with the random process that are generating the random variable based on the observed data. To predict a random variable, first we compute (in most cases) the expected value of the random variable we want to predict. Then we try to estimate that expected value based on the available information or the estimated parameters of the random process. It should be clear that the estimator itself is a function of observed data and hence a random variable and its expected value need not match the expected value of the random variable we want to predict. If the two quantities are equal, the estimate is an unbiased estimator. The unbiasedness may be desirable criteria and in many cases, it may be preferred, but it is not always a best estimate and in many cases, it may not be possible to find an unbiased estimator. If we accept this notion of estimating the parameters of the loss process, the discrepancy we observe in the number of parameters can be explained. We are estimating both the relative exposures and the payout pattern and the true number of parameters is 2(*n* - 1). The individual year ultimate losses are themselves parameters of the random process and should be counted as such when we use the weighted loss development method or the exposure development method. I would like to add one other observation that is relevant to our discussion of number of the parameters. Technically, if we are interested in total ultimate losses for all accident years combined, we need to compute just one parameter. The estimated ultimate loss for all accident year by the weighted loss development method is same as the exposure development factor $D_{n\text{-}1}$ times the first accident year total paid losses by age n. The result can also be obtained by multiplying the sum of paid losses for all accident years in the first year with the age 1 to n ultimate weighted loss development factor. This will imply that we need only one parameter in estimating the all accident years combined ultimate loss.

I would like to point out that Lehigh (2007) has expressed similar views. He states that we use losses of prior development years as a proxy for exposure. However, the fact may be that we are estimating the exposure levels as well and not realizing it.

The exposure-based method does not assume any relationship between future losses and the paid losses to date. After the Mack (1993) paper, there is strong feeling among actuaries that the use of loss development methods has an implicit assumption that future development is dependent on current observation. It was one of the basic assumptions of Mack's method that future losses depend on losses paid to date by a constant factor. Chu, and Venter (1998) discusses methods to test this assumption.

It is well known that under the assumption that $X_{i,j}$ are independently distributed Poisson or multinomial variates, the same results as the weighted loss development method are obtained and the proof can be found in Renshaw and Verrall (1998). Therefore, the claim that $2(n - 1)$ parameters are being estimated, or the losses to be paid in future are independent of paid to date, is not new. One important difference in the method presented here is that our assumptions are slightly less restrictive. Renshaw and Verral require that both the column and row sum for the observed data be positive whereas we require only row sum to be positive.

The exposure development introduced here can also use simple averages of the exposure development factors, similar to what is done in the simple average loss development method. However, the two results from loss development and exposure development will not coincide. As we shall see, in the weighted loss development method, there is a balancing going on and that causes the exposure development and loss development results to coincide. Actuaries generally prefer weighted loss development factors over simple average loss development factors. Using simple averages of the exposure development factors will be confusing if the incremental loss is negative and is therefore not recommended. However, simple averages can be used for estimating rates. It may provide an alternative estimate of the ultimate losses and can be used in making a selection of the reserve requirements. We shall return to these issues later in the paper.

In the next section, we introduce yet another alternative computational procedure that reinforces the same idea and further strengthens the view that we are estimating both exposure and payout of the ultimate losses. That computational scheme has its own merit and utility besides strengthening the ideas presented here. The computational scheme is quite versatile, and helps us in assessing the validity or the appropriateness of the model. It identifies any outliers in our data and opens up a new area for further research, as well as provides a tool for estimating the variability of our reserve estimates.

In section 3, we define the basic model of loss reserving and discuss the common thread among most of the classical actuarial methods of loss reserving. The model presented is not new and one form or another has been presented by many authors, however the perspective here is different. The reader is encouraged to read Mack and Venter to get a better understanding of the issues and controversies.

Section 4 is quite brief and focused on the basic assumptions of loss development methods and some of the actuarial adjustments that are made in practice. We also discuss the validity of the method for policy year and report year losses.

Section 5 is devoted to an example where we carry out an analysis of a selected paid loss triangle and test its appropriateness.

In section 6, we discuss variability in the estimation of ultimate losses. We provide a simple simulation approach to attack the problem but most of the details are left to the reader to extend and modify the approach as needed for analyzing the data in hand.

In section 7, we focus on the exposure development method and see how it can be used to deal with another important issue, which is using both paid and incurred loss data. As we shall see the new methodology provides us a variety of different ways to achieve it. We define several new estimators and see how information available, from incurred loss data, can be used along with paid loss data to refine our results.

## SECTION 2: INDEPENDENCE OF ACCIDENT YEAR

Most actuaries are familiar with categorical contingency tables and Chi Square test of independence. If we classify a population in two or more different categories and each of these classifications have two or more groups and we count the number of observations by category, we have a contingency table. For example, we may be interested in whether education level depends on gender. We may take a sample and count the number of people that have high school degree, a two-year college degree, a four-year college degree or a postgraduate degree separately for males and females and carry out a test to see whether education level differs for males and females. We shall not get into the computational details here, as that is not the purpose of the presentation. However, one can see the similarity and the differences with a loss triangle. The categories are accident years and development years and instead of counts we have paid loss amounts. The most important difference is that the loss dollars are not scalars and the lower half triangle of the loss rectangle is not known and our aim is to estimate them. However, it should not deter us from computing the expected value of each cell as we do in analyzing a contingency table.

Let us assume that we have all the observations in our loss rectangle. Let us define

$$R_i = \sum_{j=1}^{n} X_{i,j} \tag{2.1}$$

$$C_j = \sum_{i=1}^{n} X_{i,j}. \tag{2.2}$$

$$T = \sum_{i=1}^{n}\sum_{j=1}^{n} X_{i,j}. \tag{2.3}$$

Define

$$\hat{X}_{i,j} = R_i \times \frac{C_j}{T}. \tag{2.4}$$

However, we do not know some of the $X_{i,j}$ and aim to estimate them from the observed data to date. We shall use an iterative procedure to achieve this. We assign the value 0 to all unknown $X_{i,j}$ and use equation (2.4) to compute them. This is our first iteration and will give us an estimate of unobserved $X_{i,j}$. We substitute these estimated values in place of the previously assigned values of zero for unobserved $X_{i,j}$. We update the values of $R_i$, $C_j$, and $T$ and use equation 2.4 again to revise our estimate for unknown $X_{i,j}$. We repeat the process until it converges. The process will converge as long as each of the original $R_i$s are positive (i.e., each accident year has positive exposure). The proof is messy and left to the reader. We only state that the estimates obtained by the weighted loss development method are a solution satisfying the stated criterion. The important point to note is that the process converges to the same values as the exposure development method and the weighted loss development method. Clearly we have estimated $2(n - 1)$ parameters.

This computing method is estimating the losses to be paid for accident years 2, 3 … $n$ assuming that the loss payments are independent of accident year and that losses paid so far have nothing to do with future loss payments. A typical question one may ask is whether it is possible to test the assumption of independence. The answer is unfortunately no. One can compute statistics similar to Chi-Square as we do for contingency tables, but loss amounts are not scalar (i.e., if we restated the loss amounts in cents rather than dollars the value of the statistics so computed will be 100 times larger). We need a suitable scaling factor to test the assumption of independence. There is no satisfactory solution to the problem and we leave it as a challenge to the actuarial profession. One solution the author suggests is, if the claim count data is also available, the scaling factor can be approximated by the ratio of estimated total loss dollars for all accident years divided by the estimated total claim count for all accident years. One will divide the computed Chi-Square type statistics by this number and consider it distributed as Chi Square with $n^2 - 2n$ degrees of freedom. This technique has two problems. First, the estimated scaling factor is a random variable and second the scaling factors may be different for each cell due to inflation and varying average claim size by payment lag.

We cannot test the appropriateness of the assumption of independence of accident year and payment year lag. However, it does not prevent us from testing the suitability of the model. We have estimated both exposure and payment patterns and can obtain the estimates for each of the observed values and compute the residuals. These residuals can be tested for randomness, any pattern in accident year and payment year lag, as well as any outliers in the data. We can also compute the explained variation of the model and other statistics for goodness of fit of the model. We have analyzed a paid loss triangle data and shall discuss these results later in the paper.

One additional advantage of this iterative procedure is that we can use it when some data points are missing or when we believe the residuals are too large for some data elements and want to remove them from the analysis. These data points can be treated in the same manner as unobserved data points in the iterative estimation process. The only data elements one cannot remove are $X_{n,1}$ and $X_{1,n}$ for the obvious reasons. The removal of individual data elements and the ability to fit the original model allows us to compute model skill as introduced by Jing, Lebens, and Lowe (2009) in the actuarial field. There are additional advantages to removing a data element, as we shall see later.

## SECTION 3: BASIC MODEL OF LOSS RESERVING METHODS

We shall define a model that is basic to almost all of the classical actuarial methods.

$$X_{i,j} = a_i \times b_j + e_{ij} . \tag{3.1}$$

Where

$a_i$ is the accident year $i$ total loss,

$b_j$ is proportion of losses to be paid in payment lag $j$ and is constant for all accident years, and

$e_{ij}$ are error terms with mean zero and variance that may not be constant.

This model has $2n$ - 2 parameters, as there are 2 constraints

$$\sum_{j=1}^{n} b_j = 1 \text{ and}$$

$a_1$ is presumed known and equals $R_1$ defined earlier.

This model can be re-parameterized as

$$X_{i,j} = \mu \times a_i' \times b_j + e_{ij}, \qquad (3.2)$$

where $\mu = \sum_{i=1}^{n} a_i$ , $a_i'$ is $a_i / \mu$ and $\mu$ represents total expected loss amount for all accident years combined.

Now we shall explore the various actuarial methods and see how these are related to this basic model.

3.1 Weighted Loss Development Method: In this method the parameters of the model are estimated such that

$$\sum_{i=1}^{n+1-j} X_{i,j} = \sum_{i=1}^{n+1-j} a_i \times b_j \text{ , } j = 1, 2 \ldots n, \qquad (3.3)$$

$$\sum_{j=1}^{n+1-i} X_{i,j} = \sum_{j=1}^{n+1-i} a_i \times b_j \text{ , } i = 1, 2 \ldots n. \qquad (3.4)$$

The weighted loss development method or the exposure development method introduced here can be used to solve the above system of equations. The iterative procedure may be a systematic approach to find the same solution. We call it a systematic method merely to convey the idea that a mathematician given the problem and not exposed to actuarial methods will probably proceed that way.

3.2 Buhlman Method: We have already seen this method. In this method, *a*s are known and we estimate *b* parameters.

3.3 Bornhuetter-Ferguson Method: In this method we assume to have prior knowledge of ultimate losses. However, we do not use this information to compute the payment pattern. The payment pattern is derived as in the weighted loss development method, which presumes no knowledge of exposure or loss amounts. We then use this computed payment pattern and the prior known ultimate losses to estimate unknown loss values. The method is sometimes referred to as the combining of observed data and prior knowledge. However, this prior knowledge is not fully utilized to estimate the parameters to be used in the forecast. The method will be the same as the Buhlman method if the prior knowledge of ultimate loss is used in estimating the payment pattern.

3.4 Cape-Cod Method: This method is similar to Bornhuetter-Ferguson (B-F) method. We assume that we know the premium amount for each accident year but not the loss ratio. The loss ratio is derived from equating the actual paid to date losses for all accident years to the estimated percentage of earned premium. This method has the same basic flaw that the B-F method has. The knowledge of premium is not used in estimating the earned percentage or the payment pattern.

3.5 Least Squares Method: This method is also not that common in North America. We try to estimate $a_i$ and $b_j$s such that the residual sum of squares (RSS) is minimum, i.e.,

$$RSS = \sum_{i=1}^{n} \sum_{j=1}^{n+1-i} (X_{i,j} - a_i \times b_j)^2 \tag{3.5}$$

To solve for *a*s and *b*s, we differentiate equation 3.5 with respect to $a_i$s and $b_j$s and equate them to zero. The derived set of equations requires an iterative procedure for solution. We shall not pursue it here. A variation of this method is to weigh the individual error term by some predefined weighting factors.

3.6 Log Regression Model: This is a new trend in the last few decades but it is still not widely used in practice. The basic model is the same as equation (3.2) with one basic difference. The error terms are assumed to be multiplicative and have mean 1 rather than additive with mean 0. One takes the logarithm of the paid incremental losses, and the model becomes linear in parameter. These new parameters can be estimated much more easily. Interested readers are referred to Verral (1994). The modeling process breaks down if some of the paid values are negative and a variety of ad hoc adjustments are made to the data are made to fit the model and estimate the model parameters and the unpaid losses. The main drawback of this method is that it requires transforming the data by taking logarithms. Once we have estimated the parameters we have to convert the estimates to original units. There are many advantages as we can test the significance of the various parameters and can define the parameters in some functional form and reduce the number of the parameters to be estimated. The transformed equation (3.2) can be modified to include the calendar year parameters. There is vast literature on this methodology and we will not pursue it here. Alternative transformations other than logarithmic are also investigated by a few authors.

It may be worthwhile to add that the iterative procedure introduced in section 2 provides many of the advantages of this methodology. In section 5 we have a numerical example and discuss it in detail.

## SECTION 4: INFLATION EFFECT

We have seen that for most of the actuarial methods, the basic underlying model is the same. In this section, we discuss the effect of inflation on the basic model as well as some of the simple approaches used by actuaries to deal with it.

The basic model presumes that each accident year has an exposure level (ultimate losses); losses are paid by a fixed pattern and that pattern remains constant over time. These are the implications of the assumption that the claims reporting and handling process is same for all accident years. Any changes we may observe are due to randomness and not due to systematic changes in the loss process or claims handling. We know that inflationary changes affect the loss payments. Under the assumption that inflation affects the loss payment by accident year only, the basic model is not affected. Inflation affects the losses paid uniformly for each delay and the payment pattern will remain the same for all accident years. The inflation impact will be in parameters $a_i$s only and will be captured by the estimation process. However, the losses paid may be impacted by both the accident year as well as the year losses are paid. Bustic (1988) discusses these issues in detail. Under this scenario, the payment pattern is affected and the model (3.1) is distorted. The best way to handle such a situation is to restate the loss triangle by removing the inflationary effect, estimate the parameters, and adjust the estimated losses for the inflation. However, this may add more estimation error in our analysis. First, we have to estimate the inflation by accident year and how the loss payment is affected by payment delay and the accident year. There is no simple solution to these estimations, thus adjusting the loss triangle for inflation may add more distortion in the results rather than improving it. One common technique used by most actuaries is to compute the loss development factors based on more recent data (latest three years' average development factors). If we assume that either inflation changes for each year but changes are moderate or the effect of the payment lag is small or both, this adjustment works well. One of the advantages of the approach that we estimate both exposure level and the payment pattern is that the use of the latest years in estimating parameters can be modified. We could use it for exposures only or rates or both and as such providing us with alternative estimators. The concept is made clearer when we analyze a loss triangle later in the paper.

The assumption that we are estimating both the exposure level and the payment patterns raises another issue of great importance. Actuarial literature encourages the use of the loss development method for policy year loss triangles as well as report year loss triangles. Under the assumption that the exposure level is also being estimated, the loss development methodology is inappropriate for analyzing report year loss triangles. Each element of a report year loss triangle will have losses generated from a different number of accident years and the exposure level keeps changing for such a loss triangle. For policy year loss triangles, the inflationary changes will distort the data much more severely as they are affected by two years of inflationary impact. Unless inflation is fairly constant, the use of exposure development method on a policy year loss triangle may be questionable. However, it will lead to the same result as the weighted loss development method and indirectly raises questions about the suitability of using the loss development method for the policy year loss triangle. The inflationary distortion will be much more significant in a policy year loss triangle if the inflationary changes are large. Although, this author has no serious objection to the use of loss development method to the policy year loss triangle, however the additional analysis carried out in the next section, especially the testing the model validity and defining outliers, may not be appropriate for such data. We have also provided a method for computing variability in the loss reserve. Such an analysis for policy year loss triangles may be distorted.

## SECTION 5: NUMERICAL EXAMPLE

We now focus on analyzing a real data set. This will help create a clearer understanding of the ideas presented in this paper.

We have selected a data set for use in this example; the main reason for selecting this data was that both the paid and incurred loss triangles are available. We can see how the information from both triangles is combined to estimate ultimate losses. In this section we focus on paid losses only.

We shall use model (3.2) for our discussion. We use a paid loss triangle from Quarg and Mack (2008) that has seven years of data. The incremental paid loss triangle, the development factors, and some additional computations are given below in table 1.

**Table 1**

| | Payment Lag | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **AY** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | $R_i$ | **Ultimate** | $a_i$ |
| 1 | 576 | 1228 | 166 | 54 | 50 | 28 | 29 | 2131 | 2,131 | 0.0677 |
| 2 | 866 | 1082 | 214 | 70 | 52 | 64 | | 2348 | 2,380 | 0.0757 |
| 3 | 1412 | 2346 | 494 | 164 | 78 | | | 4494 | 4,652 | 0.1479 |
| 4 | 2286 | 3006 | 432 | 126 | | | | 5850 | 6,182 | 0.1965 |
| 5 | 1868 | 1910 | 870 | | | | | 4648 | 5,056 | 0.1607 |
| 6 | 1442 | 2568 | | | | | | 4010 | 4,934 | 0.1568 |
| 7 | 2044 | | | | | | | 2044 | 6,128 | 0.1948 |
| Total | 10,494 | 12,140 | 2,176 | 414 | 180 | 92 | 29 | **25,525** | **31,463** | |
| DF | 2.437 | 1.131 | 1.029 | 1.021 | 1.021 | 1.014 | | | | |
| CDF | 2.998 | 1.23 | 1.088 | 1.057 | 1.035 | 1.014 | 1 | | | |
| $b_j$ | 0.334 | 0.479 | 0.106 | 0.027 | 0.02 | 0.02 | 0.014 | | | |
| $\mu \times bj$ | **10,494** | **15,077** | **3,356** | **849** | **618** | **642** | **428** | 31,463 | | |

For simplicity, we have computed ultimate losses using the loss development method. They could have easily been computed using an iterative procedure. The column $\mathbf{a}_i$ is accident year ultimate losses divided by the sum of estimated ultimate losses for all accident years, and represents the proportion of total losses for the accident year. We shall use the term exposure level to represent this quantity. The bottom two rows are the payment pattern and the total losses for the payment lag respectively. If we used the iteration procedure, the solution would converge at these values. In table 2 below, we give the residuals for each accident year and payment year. These are computed by subtracting the estimated values from observed data. The estimated values are the bottom row times the $\mathbf{a}_i$s for the corresponding row and columns.

**Table 2**

| **Residuals** | Payment Year | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| AY | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Sum |
| 1 | (134.76) | 206.86 | (61.30) | (3.49) | 8.14 | (15.46) | - | 0.00 |
| 2 | 72.06 | (58.64) | (39.90) | 5.78 | 5.24 | 15.46 | | 0.00 |
| 3 | (139.65) | 116.76 | (2.21) | 38.49 | (13.38) | | | 0.00 |
| 4 | 224.23 | 43.89 | (227.35) | (40.77) | | | | 0.00 |
| 5 | 181.79 | (512.55) | 330.76 | | | | | 0.00 |
| 6 | (203.68) | 203.68 | | | | | | 0.00 |
| 7 | 0.00 | | | | | | | 0.00 |
| Sum | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | **0.00** |

Looking at these residuals, the second payment for accident year 5 seems to be an outlier. One can remove this observation and revise the estimate. We will be constructing this estimate later in the paper for estimating the variability of our reserve estimates. The residuals can be further analyzed as to whether there is a systematic variation from the model and some adjustments to the model can be made as needed. For the current data set the model seems quite good. The model statistics are given in the table below in table 3.

**Table 3**

| | Sum of Squares | DF | MSS |
|---|---|---|---|
| Total | 23,568,917 | 27 | 872,923 |
| Error | 704,033 | 12 | 58,669 |
| Explained | 22,864,884 | 15 | 1,429,055 |
| | | F | 24.36 |
| | | $R^2$ | 0.97 |

The $R^2$ is unusually high for this data set and tells us that the estimated parameters fit the model very well. We have computed some basic model testing statistics. One may compute a host of other statistics for testing the appropriateness of the model. We shall not pursue these in detail, as that is not the theme of the paper. We shall focus on skill of the model statistics recently introduced by Yi Jing, Joseph R. Lebens, and Stephen P. Lowe (2009) to the actuarial field. However, they used it quite differently by computing it through the observed future with predicted future. The modeling procedure presented here allows us to compute it for a current data set and test how good the model will be for predicting the future. It may be a bit confusing that we need to look for additional statistics even if the explained ratio is quite high or other statistics indicate that the model is a good fit. One can think of the skill of the model as testing for model specification error. The assumption that we estimated both the exposure level as well as the payment pattern allows us to estimate the model skill. We have mentioned before that the iterative procedure can be used by removing individual observations. The skill of a model is defined as

$$Skill = 1 - \frac{SSA}{SSE} \tag{5.1}$$

where SSE is the average squared error of estimation by fitting all observed data points,

and SSA is the average squared error of estimation error of individual observations estimated by removing that observation and estimating it from the remaining observations. This following example will help clarify. We remove the first observed value from our data set and estimate the parameters. These parameters provide a new estimate for $X_{11}$. The original estimate of $X_{11}$ was obtained by using all data points including observed $X_{11}$. We do this for each of the other observations. The square of the error of the second estimate from the observed value is averaged over all data points to compute SSA. In our case we can compute it for all but two observations. The following table displays the results of this computation along with some additional data that we will need for analysis in the next section.

**Table 4**

| | | | | Exposure Proportion | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| AY | Lag | ESS | Error | AY 1 | AY 2 | AY 3 | AY 4 | AY 5 | AY 6 | AY 7 |
| 1 | 1 | 697,832 | 226.5 | 0.0749 | 0.0755 | 0.1474 | 0.1957 | 0.16 | 0.1559 | 0.1907 |
| 1 | 2 | 818,383 | -438 | 0.0544 | 0.0767 | 0.1503 | 0.1999 | 0.1637 | 0.1602 | 0.1948 |
| 1 | 3 | 698,045 | 77 | 0.07 | 0.0754 | 0.1473 | 0.1957 | 0.16 | 0.1568 | 0.1948 |
| 1 | 4 | 703,998 | 4 | 0.0678 | 0.0756 | 0.1478 | 0.1964 | 0.1607 | 0.1568 | 0.1948 |
| 1 | 5 | 704,770 | -11 | 0.0675 | 0.0757 | 0.148 | 0.1965 | 0.1607 | 0.1568 | 0.1948 |
| 1 | 6 | 702,858 | 30 | 0.0682 | 0.0751 | 0.1479 | 0.1965 | 0.1607 | 0.1568 | 0.1948 |
| 2 | 1 | 715,150 | -121 | 0.0678 | 0.0718 | 0.1481 | 0.1969 | 0.161 | 0.1573 | 0.197 |
| 2 | 2 | 718,065 | 129 | 0.0674 | 0.0795 | 0.1472 | 0.1955 | 0.1598 | 0.1558 | 0.1948 |
| 2 | 3 | 704,950 | 51 | 0.0676 | 0.0771 | 0.1475 | 0.196 | 0.1602 | 0.1568 | 0.1948 |
| 2 | 4 | 703,358 | -7 | 0.0678 | 0.0755 | 0.1479 | 0.1966 | 0.1607 | 0.1568 | 0.1948 |
| 2 | 5 | 703,700 | -7 | 0.0678 | 0.0755 | 0.148 | 0.1965 | 0.1607 | 0.1568 | 0.1948 |
| 2 | 6 | 703,051 | -33 | 0.0682 | 0.0751 | 0.1479 | 0.1965 | 0.1607 | 0.1568 | 0.1948 |
| 3 | 1 | 723,121 | 267 | 0.0675 | 0.0754 | 0.156 | 0.1956 | 0.1599 | 0.1557 | 0.1899 |
| 3 | 2 | 740,956 | -286 | 0.0685 | 0.0765 | 0.1399 | 0.1987 | 0.1626 | 0.159 | 0.1948 |
| 3 | 3 | 703,886 | 3 | 0.0677 | 0.0756 | 0.1479 | 0.1964 | 0.1607 | 0.1568 | 0.1948 |
| 3 | 4 | 702,123 | -57 | 0.068 | 0.0759 | 0.1466 | 0.1972 | 0.1607 | 0.1568 | 0.1948 |
| 3 | 5 | 704,873 | 28 | 0.0675 | 0.0754 | 0.1483 | 0.1965 | 0.1607 | 0.1568 | 0.1948 |
| 4 | 1 | 894,943 | -453 | 0.0682 | 0.0762 | 0.1489 | 0.1825 | 0.162 | 0.1586 | 0.2036 |
| 4 | 2 | 725,306 | -119 | 0.0681 | 0.076 | 0.1486 | 0.1934 | 0.1615 | 0.1577 | 0.1948 |
| 4 | 3 | 712,011 | 375 | 0.0665 | 0.0742 | 0.1451 | 0.2051 | 0.1575 | 0.1568 | 0.1948 |
| 4 | 4 | 702,149 | 71 | 0.0674 | 0.0753 | 0.1471 | 0.1979 | 0.1607 | 0.1568 | 0.1948 |
| 5 | 1 | 667,123 | -352 | 0.0682 | 0.0761 | 0.1488 | 0.1977 | 0.1494 | 0.1582 | 0.2015 |
| 5 | 2 | 2,404,396 | 1436 | 0.0639 | 0.0714 | 0.1395 | 0.1854 | 0.1984 | 0.1466 | 0.1948 |
| 5 | 3 | 674,610 | -485 | 0.0695 | 0.0777 | 0.1518 | 0.2017 | 0.1477 | 0.1568 | 0.1948 |
| 6 | 1 | 771,868 | 438 | 0.067 | 0.0748 | 0.1462 | 0.1942 | 0.1589 | 0.172 | 0.187 |
| 6 | 2 | 938,021 | -598 | 0.0698 | 0.0779 | 0.1523 | 0.2024 | 0.1655 | 0.1374 | 0.1948 |

The first two columns represent the accident year and the payment year of the observation that was removed from the estimation process. The third column is the total error sum of squares for all observed values and column four is the estimation error of the observed value that was removed from the fitting. One can see that the error sum of squares are comparable to the error sum of squares of 704.03, which was computed based on fitting the model to all data points except for the error sum of squares for the second payment for accident year 5. Most of this variation is coming from the estimation error of this observation itself, as the corresponding residual is quite high (1,436 in the table). This observation is over-estimated a little more when it is removed from the fitting. This gives further credence to the previous statement that this observed value is probably an outlier in the data set. The data set overall appears to be well-behaved and the model appears to perform quite well as the total error sum of squares remains fairly constant when other individual data points are removed from the estimation process. We also captured the estimated accident year contribution to the all accident year estimated ultimate loss in each scenario, which we shall be using in estimating variance. These values are in columns 5 to 11.

The skill of the model is one minus the average of sum of squares of column 4 divided by the average error sum of squares with all data points included in the analysis. Its value is 0.79 for this data.

We will not pursue here the removal of the outliers and revising the estimates. We only broach this issue to point out that the modeling process presented allows us to identify such data elements and adjustments can be made as warranted. However, removal of the second payment for accident year 5 will result in accident year 5 ultimate losses of 6,617 instead of 5,056.

In table 5, we provide our analysis for the corresponding incurred loss triangle.

**Table 5**

| | Payment Lag | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **AY** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | **Sum** | **Ult** | $a_i$ |
| 1 | 978 | 1126 | 30 | 10 | 30 | 8 | -8 | 2174 | 2,174 | |
| 2 | 1844 | 708 | -86 | 14 | 28 | -54 | | 2454 | 2,445 | 0.0739 |
| 3 | 2904 | 1450 | 344 | -98 | 44 | | | 4644 | 4,582 | 0.1385 |
| 4 | 3502 | 2456 | 112 | 72 | | | | 6142 | 6,126 | 0.1852 |
| 5 | 2812 | 2070 | -30 | | | | | 4852 | 4,839 | 0.1463 |
| 6 | 2642 | 1764 | | | | | | 4406 | 4,476 | 0.1353 |
| 7 | 5022 | | | | | | | 5022 | 8,429 | 0.2549 |
| Sum | 19,704 | 9,574 | 370 | -2 | 102 | -46 | -8 | **29,694** | **33,071** | |
| DF | 1.65 | 1.02 | 1 | 1.01 | 0.99 | 1 | | | | |
| CDF | 1.68 | 1.02 | 1 | 1 | 0.99 | 1 | 1 | | | |
| $b_j$ | 0.596 | 0.389 | 0.018 | 0.000 | 0.011 | -0.010 | -0.004 | | | |
| $\mu \times bj$ | **19,704** | **12,849** | **607** | **-4** | **367** | **-329** | **-122** | 33,071 | | |

The estimated ultimate losses from the incurred loss triangle are higher than the paid loss triangle. Accident year 7 is contributing for most of this difference. There is a significant increase in first year incurred loss for accident year 7 compared to earlier accident years. The paid loss triangle does not show such an increase. One will probably give less credence to the ultimate losses derived from incurred loss triangle for accident year 7 unless there is significant increase in the volume of business and is known from some alternative sources.

## SECTION 6: VARIABILITY IN LOSS RESERVES

The estimation of variability in loss reserves is becoming an important issue. Although there are some methods available to achieve this, there is no consensus in the actuarial profession. Ad hoc methods are commonly used to derive a range of estimates. One uses a variety of methods or a different data set, paid and incurred loss triangles for example, to derive a range for ultimate losses. A range for ultimate losses is achieved but the assigning of a confidence level is not possible when these types of methods are used. We shall develop a simulation methodology to estimate the variability of the reserve estimates.

We shall again assume that the exposure levels are known and compute its variability. We shall use model (3.2) and further assume that

$$V(e_{ij}) = a_i \times \sigma_j^2 . \tag{6.1}$$

Under these assumptions

$$\hat{\mu} \times b_j = \frac{\sum_{i=1}^{n-j+1} X_{i,j}}{\sum_{i=1}^{n-j+1} a_i}, \tag{6.2}$$

$$\hat{\sigma}_j^2 = \left( \sum_{i=1}^{n-j+1} \frac{X_{i,j}^2}{a_i} - \frac{\sum_{i=1}^{n-j+1} X_{i,j}^2}{\sum_{i=1}^{n-j+1} a_i} \right) \times \frac{1}{n-j+1} \tag{6.3}$$

Since we have only one observation for payment year $n$, the variance cannot be estimated for that period. For our computational example, we have estimated the variance for $b_n$ by the maximum of the variance estimates of $b_{n-1}$ and the average of the variance estimates of $b_{n-1}$ and $b_{n-2}$.

It must be noted that the variance assumption in equation (6.1) may not be valid. Exposure changes are caused by two factors: changes in volume cause the variance to increase linearly, which is consistent with equation (6.1), and changes in inflation cause variance to increase exponentially. Our formulation of the model is consistent with the way parameters are being estimated. Large changes in inflation may cause this variance to be underestimated slightly.

Under the assumption of independence of future payments,

$$\hat{X}_{i,j} = a_i \times \hat{\mu} \times b_j, \tag{6.4}$$

$$\hat{V}(X_{i,j}) = a_i \times \hat{\sigma}_j^2 \times \left(1 + a_i \times \left(\sum_{i=1}^{n-j+1} a_i\right)^{-1}\right). \tag{6.5}$$

However, $a_i$s are not known and are estimated from the same data. Hence our estimate of the variance is understated. We will attack this problem by using bootstrap and simulation methods and use the following well-known equation. It is worth mentioning that equation (6.5) defines the variance for individual incremental payments. The all accident year variance estimates will be larger than the sum of individual accident years due to correlation introduced in accident year estimates by the estimation process.

$$V(X) = E\left(V\left(\frac{X}{A}\right)\right) + V\left(E\left(\frac{X}{A}\right)\right). \tag{6.6}$$

In the previous section we computed values $a_i$ by reducing our observation set by one observation at a time. We can use the results for the exposure levels captured there for estimating the variance of the estimation through simulation. Steps of our simulation approach are as follows.

Step 1. Find minimum and maximum values for each accident year for columns 5 to 11 from table 4.

Step 2. Generate a uniform random variable in the range between minimum and maximum values for each accident year. These are preliminary relative exposures for each of the accident years.

Step 3. These exposure levels will not add to 1. Normalize them by dividing each preliminary exposure by the sum of the preliminary exposure levels.

Step 4. Use the normalized exposure levels in equation (6.2) to (6.5) to estimate the $X_{i,j}$ and its variance.

Step5. Repeat the process 1,000 times and use these to estimate the terms in equation (6.6); treat the result of each iteration as an observation of the corresponding variable.

One can increase the number of iterations if the data has larger variation. One thousand iterations for the current data set were sufficient.

The results for the paid loss triangle are summarized below for each accident year as well as totals for all accident years. One should note that the variance for all accident years is larger than the sum of individual accident years.

**Table 6**

**Simulation Results**

| AY | Ultimate | Sim-d Ave Ultimate | Expected Variance | Variance Expected | Estimated Variance | ST Dev |
|---|---|---|---|---|---|---|
| 1 | 2,131 | 2,131 | 0 | 0 | - | - |
| 2 | 2,380 | 2,382 | 628 | 11 | 639 | 25 |
| 3 | 4,652 | 4,659 | 3,677 | 125 | 3,802 | 62 |
| 4 | 6,182 | 6,188 | 6,310 | 373 | 6,683 | 82 |
| 5 | 5,056 | 5,097 | 6,704 | 1,672 | 8,376 | 92 |
| 6 | 4,934 | 4,928 | 45,332 | 4,091 | 49,423 | 222 |
| 7 | 6,128 | 6,131 | 295,912 | 20,055 | 315,967 | 562 |
| **Total** | **31,463** | **31,515** | **450,987** | **26,873** | **477,860** | **691** |

The all year total variation is larger than sum of individual accident years because of correlation.

## SECTION 7: EXPOSURE DEVELOPMENT METHOD

The concept of the exposure development factor (EDF) method introduced in this paper is very useful. One important area where a lot of attention is being paid is combining the information from paid and incurred loss triangles to refine our estimates. In the 2009 CLRS meeting, there was a full session devoted to this topic. The EDF method provides an elegant way to achieve this. The important characteristic of the EDF method is that, unlike loss development factors, the EDFs for paid and incurred loss triangles are measuring the same quantity and provide two estimates of the relative exposure levels. This property can be exploited with significant improvement in our analysis of loss triangles. One extreme will be to use exposure levels derived from the paid loss triangle to the incurred loss triangle and vice versa. A better way would be to average the exposure levels determined by the paid and incurred loss triangles. The exposure levels from two triangles will be correlated, as the paid losses are included in the incurred losses. The average of the two factors will still be a better estimate. The averaging can be done in a variety of ways. One can average the year-to-year exposure development factors or the normalized exposure levels. One could use differential weights as well.

Once the selection of exposure level for each accident year is made, we use it to determine the payout pattern. In the examples presented earlier, we have used combined payout for all years. However, one can determine each accident year's payout rate separately and then make a selection.

In the loss development method, actuaries use a variety of averaging procedures and professional judgment to select a development factor. Similar analysis can be carried out in determining rates for the selected exposure level. One can take an average after removing high and low values for rates, for example.

In the following table we provide an example. The main purpose of this is to show how the data from the different triangles can be combined and used in a systematic way. In the table below we have adopted an arbitrary weighting scheme to select accident year exposure levels.

**Table 7**

| AY | Paid Exposure Level | Incurred Exposure Level | Weight | Weighted Exposure Level | Selected Exposure Level |
|---|---|---|---|---|---|
| 1 | 0.0677 | 0.0657 | 0.5 | 0.0667 | 0.0684 |
| 2 | 0.0757 | 0.0739 | 0.5 | 0.0748 | 0.0766 |
| 3 | 0.1479 | 0.1385 | 0.5 | 0.1432 | 0.1467 |
| 4 | 0.1965 | 0.1852 | 0.5 | 0.1909 | 0.1955 |
| 5 | 0.1607 | 0.1463 | 0.25 | 0.1499 | 0.1536 |
| 6 | 0.1568 | 0.1353 | 0.25 | 0.1407 | 0.1442 |
| 7 | 0.1948 | 0.2549 | 0.75 | 0.2098 | 0.215 |
| **Total** | **1.0000** | **1.0000** | | **0.9760** | **1.0000** |

We have changed weights for accident year 5, 6, and 7. We saw before that the second payment for accident year 5 might be an outlier. It will affect EDFs 4 and 5 and exposure levels so less weight is assigned to the exposure level derived from the paid triangle for these years. The incurred loses for accident year 7 is quite high compared to accident year 6. We do not see that magnitude of increase in paid losses. More weight is therefore given to the exposure level derived from the paid loss triangle.

Now we use these selected exposure levels and the total observed payout by delay for each accident year and select a payout judgmentally. We are a bit conservative in our selection. This is obvious from the fact that the total estimated payout is less than the selected payout.

**Table 8**

| AY | Delay1 | Delay2 | Delay3 | Delay4 | Delay5 | Delay6 | Delay7 | Exposure Level |
|---|---|---|---|---|---|---|---|---|
| 1 | 576 | 1,228 | 166 | 54 | 50 | 28 | 29 | 0.0684 |
| 2 | 866 | 1,082 | 214 | 70 | 52 | 64 | | 0.0766 |
| 3 | 1,412 | 2,346 | 494 | 164 | 78 | | | 0.1467 |
| 4 | 2,286 | 3,006 | 432 | 126 | | | | 0.1955 |
| 5 | 1,868 | 1,910 | 870 | | | | | 0.1536 |
| 6 | 1,442 | 2,568 | | | | | | 0.1442 |
| 7 | 2,044 | | | | | | | 0.2150 |
| **Sum** | 10,494 | 12,140 | 2,176 | 414 | 180 | 92 | 29 | 1.0000 |

**Accident Year Payout**

| **Payout** | Delay1 | Delay2 | Delay3 | Delay4 | Delay5 | Delay6 | Delay7 | |
|---|---|---|---|---|---|---|---|---|
| 1 | 8,424 | 17,960 | 2,428 | 790 | 731 | 410 | 424 | |
| 2 | 11,301 | 14,119 | 2,793 | 913 | 679 | 835 | | |
| 3 | 9,624 | 15,990 | 3,367 | 1,118 | 532 | | | |
| 4 | 11,690 | 15,372 | 2,209 | 644 | | | | |
| 5 | 12,162 | 12,435 | 5,664 | | | | | |
| 6 | 10,002 | 17,812 | | | | | | |
| 7 | 9,509 | | | | | | | |
| **Average** | 10,387 | 15,615 | 3,292 | 866 | 647 | 622 | 424 | **31,854** |
| **Weighted** | 10,494 | 15,464 | 3,395 | 850 | 617 | 634 | 424 | **31,879** |
| **Selected** | **10,450** | **15,600** | **3,350** | **860** | **640** | **630** | **424** | **31,954** |

**Estimated Payout all Accident Years**

| **AY** | Delay1 | Delay2 | Delay3 | Delay4 | Delay5 | Delay6 | Delay7 | **Ultimate** |
|---|---|---|---|---|---|---|---|---|
| **1** | 576 | 1,228 | 166 | 54 | 50 | 28 | 29 | 2,131 |
| | 866 | 1,082 | 214 | 70 | 52 | 64 | 32 | 2,380 |
| | 1,412 | 2,346 | 494 | 164 | 78 | 92 | 62 | 4,649 |
| | 2,286 | 3,006 | 432 | 126 | 125 | 123 | 83 | 6,181 |
| | 1,868 | 1,910 | 870 | 132 | 98 | 97 | 65 | 5,040 |
| | 1,442 | 2,568 | 483 | 124 | 92 | 91 | 61 | 4,861 |
| | 2,044 | 3,353 | 720 | 185 | 138 | 135 | 91 | 6,666 |
| **Sum Proj** | 10,494 | 15,493 | 3,379 | 855 | 633 | 631 | 424 | **31,909** |

The incurred loss triangle can be analyzed similarly using the selected exposure levels. We shall not do it here.

Actuaries often use recent accident year data for loss development factor calculations and projections of ultimate losses. Such results are responsive to changes that are too complex to model. The exposure development method is much more flexible and therefore can achieve this. Some care is needed, as the loss payment amount in later lags may be quite thin. It is advisable to use all payment lag data of an accident year for computing the exposure development factors. In the example below, we use the available latest three accident years to compute our exposure development factors. One can directly use these development factors to compute ultimate losses.

However, we have computed payout rates as there is flexibility here. One can use all years' data or the latest three years to determine rates. If we use the latest three years' data, the results will match with the latest three-year weighted loss development method.

One alternative approach that this author prefers is to use all accident year data for exposure development factors and use the latest years' observations for selecting payout rates. Of course, one would use exposure levels derived from incurred loss triangles if available, and compute payout rates based on the latest years or by excluding Hi-Low rates as is done in selecting development factors.

One other possible variation is indicated by examining the incurred loss triangle. The incremental incurred losses for some accident years are negative possibly due to some recoveries or subrogation. These just add additional variation in EDFs. One could compute the EDFs without these values. These data points could be included in computing rates.

## SECTION 8: CONCLUSION AND FUTURE RESEARCH

In this paper we have a methodology that in some sense diverges from the common way actuaries look at loss triangles. Results are, however, consistent with loss development method and extend it in several ways. In practice, actuaries use a lot of professional judgment. Allowing judgment to be applied to both the exposure level and payment pattern, we have a two-dimensional selection processes rather than one. Knowledge of both the paid and incurred loss triangles extends that even further. The fact that the EDF method measures the same thing for paid and incurred losses has one other nice implication for excess and reinsurance writers. The paid loss experience is thin and not credible in the first few years. However, the exposure levels derived from incurred loss triangles for early years can be used on paid loss data. We had avoided the issue of tail losses. Perhaps one can use both the paid and incurred rates to derive a suitable decay function.

The author believes that the ideas presented will stimulate other researchers to modify and extend it further. There is ample opportunity to do so. We defined a range of exposure levels by removing one observation at a time and re-computing exposure levels. There may be different ways to achieve this result. One may define a range based on paid and incurred loss triangles or use information from both data sets or premium data. The simulation results in our example assumed uniform distribution in the range. One could use alternative distributions somehow derived from the data. Uniform distributions increase the variance estimates and, in that sense, are conservative estimates of the variance. Estimation of tail factors is another area where further research will be helpful.

The methodology presented in this paper is simple and is for practical use. How it fares in practice can only be determined by practicing actuaries.

**Acknowledgements**

Author wishes to thank James Heer and Manisha Srivastava for many helpful comments that significantly improved the quality of the presentation.

## REFERENCES

[1.] Butsic, Robert, "The Effect of Inflation on Losses and Premiums for Property- Liability Insurers," Inflation Implications for Property-Casualty Insurance, Casualty Actuarial Society *Discussion Paper Program*, 1981, pp. 58–102.

[2.] Chu, Julia Feng-Ming, Gary G. Venter, "Testing the Assumptions of Age-To-Age Factors," *Proceedings of the Casualty Actuarial Society Casualty Actuarial Society* LXXXV, 1998, pp. 807-847.

[3.] Halliwell, Leigh Joseph, "Chain-Ladder Bias: Its Reason and Meaning," *Variance* 1:2, 2007, pp.214-247, http://www.variancejournal.org/issues/01-02/214.pdf.

[4.] Jing, Yi, Joseph R. Lebens, and Stephen P. Lowe, "Claim Reserving: Performance Testing and the Control Cycle," *Variance* 3:2, 2009, pp. 161-193, http://www.variancejournal.org/issues/03-02/161.pdf.

[5.] Quarg, Gerhard and Thomas Mack, "Munich Chain Ladder: A Reserving Method that Reduces the Gap between IBNR Projections Based on Paid Losses and IBNR Projections Based on Incurred Losses," Variance 2:2, 2008, pp. 266-299, http://www.variancejournal.org/issues/02-02/266.pdf.

[6.] Mack, Thomas, "Distribution-free Calculation of the Standard Errors of Chain Ladder Reserves," *ASTIN Bulletin*: 23:2, 1993, pp. 213-225.

[7.] Mack, Thomas, Gary G. Venter, "A Comparison of Stochastic Models that Reproduce Chain Ladder Reserve Estimates," *Insurance: Mathematics and Economics,* Vol. 26, Issue 1, 1 February 2000, pp. 101-107.

[8.] Verrall, Richard, "Statistical Methods for the Chain Ladder Technique" Casualty Actuarial Society *Forum* Spring 1994, Vol. 1, pp. 393-446.

[9.] Renshaw, A.E., R.J. Verrall, "A Stochastic Model Underlying the Chain-Ladder Technique," *British Actuarial Journal* 4:4, 1998, pp. 903-923.