An Introduction to Credibility

by Curtis Gary Dean, FCAS

This paper is derived from the presentation on basic credibility concepts that the author has given at the 1995 and 1996 CAS Seminars on Ratemaking.

AN INTRODUCTION TO CREDIBILITY

Credibility theory provides important tools to help the actuary deal with the randomness inherent in the data that he or she analyzes. Actuaries use past data to predict what can be expected in the future, but the data usually arises from a random process. In insurance, the loss process that generates claims is random. Both the number of claims and the size of individual claims can be expected to vary from one time period to another. If $1,500,000 in losses were paid by an insurer during the past year, one might estimate that $1,500,000 would likely be paid in the current year for the same group of policies. However, the expected accuracy of the estimate is a function of the variability in losses. Using credibility theory, the actuary estimates the randomness inherent in the data and then calculates a numeric weight to assign to the data.

Here is a dictionary definition of credible:

<u>credible</u>: Offering reasonable grounds for being believed

The actuary wants to know how much to believe the data that's being analyzed. To use the data to predict the future, this "belief in the data" must be quantified so that calculations can be made. This leads us to actuarial credibility:

<u>actuarial credibility</u>: the weight to be given to data relative to the weight to be given to other data

If we cannot fully believe our data, we may call on other information or data to supplement the data at hand. The data at hand and the supplemental data are each given an appropriate numeric weight in calculating an estimate.

The variability in insurance loss data can be seen in Table 1 which shows the loss experience for a group of policies covering contractor's pickup trucks. The last column shows that the average loss per truck varies widely from one year to the next. Any one year is a poor predictor of subsequent years.

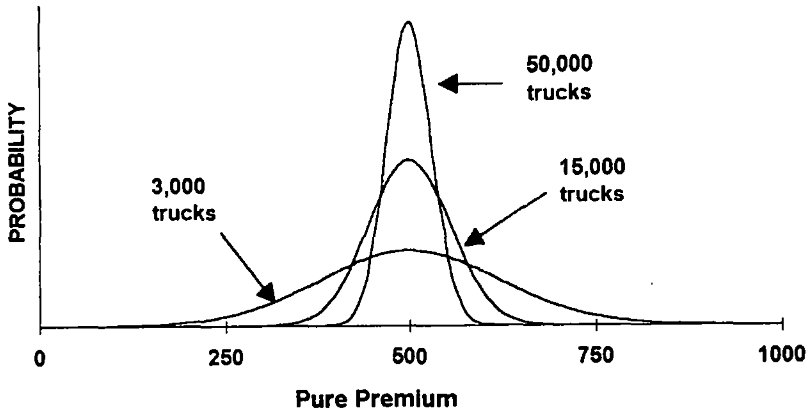
TABLE 1

Contractor's Pickup Trucks

	(1) # of Insured Trucks	(2) Incurred Losses	Pure Premium (2)/(1)
1990	2,900	$2,030,000	$700
1991	3,000	1,470,000	490
1992	3,050	1,830,000	600
1993	3,050	1,250,500	410
1994	3,200	864,000	270

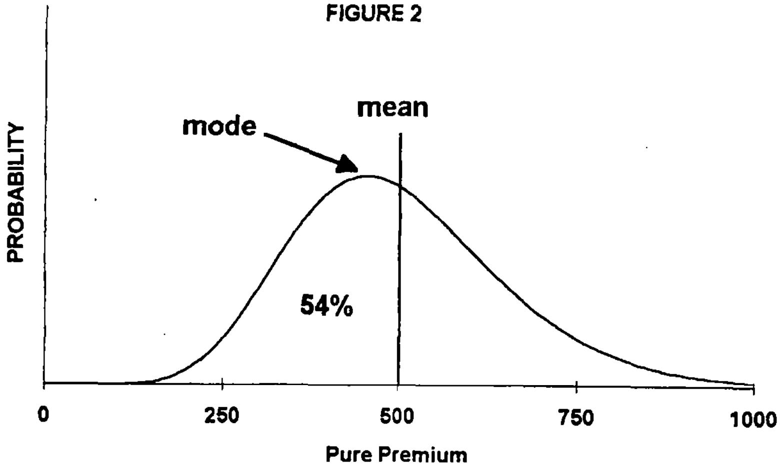
The variability in the average loss per pickup truck is depicted graphically in Figure 1. The expected average loss (pure premium) is $500 which we would observe if our body of data were infinite in size. But, for limited sample sizes, the observed average losses are randomly distributed. Note that as our sample size increases, the variability of the observed average loss decreases - the probability density curve becomes more concentrated around the $500 value. For a smaller sample size, the probability density curve flattens out. If our sample body of data consists of 50,000 trucks we can rely upon the observed average loss to estimate the true expected average loss to a much greater extent than if the data came from a smaller sample of only 3000 trucks.

FIGURE 1



The actual distribution of pure premiums is not symmetric as shown in the prior graph, but is instead skewed to the right as shown in Figure 2. More of the observations would actually fall below the mean of $500 and the mode of the distribution is less than $500. The smaller the body of data, the greater the asymmetry in the graph. In an extreme case we could consider only one truck. In most years the truck would have no losses for an observed average loss of $0 in those loss-free years. But, every few years there would be a loss or, perhaps, several losses and the observed average loss would be substantial.

FIGURE 2



This leads us to a common problem that may occur when a group of non-actuaries is reviewing average losses or loss ratios for a series of years. The data may show, for example, four years with excellent loss ratios but a fifth year with a very high loss ratio. The five-year average may be close to some target loss ratio. Unfortunately, what frequently happens is that one of the reviewers will say that the one bad year is an anomaly that was caused by several severe claims and that the bad year should be thrown out of the data. This is a big mistake! For a small body of data, this pattern in the loss ratios is exactly what we expect to see. The majority of the loss ratios will look better than average, with a few being quite large. This doesn't mean that we should ignore the few high values; it usually means that our body of data is small.

The basic formula for calculating credibility weighted estimates is:

$$\text{Estimate} = Z \times [\text{Observation}] + (1-Z) \times [\text{Other Information}],$$

$$\text{and } 0 \leq Z \leq 1.$$

If our body of data is so large that we can give full weight to it in making our estimate, then we would set Z=1. If the data is not fully credible, then Z would be a number somewhere between 0 and 1. What is the "Other Information" that we might use in our formula? That depends on what we are trying to estimate. In Table 2, the left hand column shows our observed data and the right hand column may be the "Other Information" that we might use in the above formula.

TABLE 2

Observation		**Other Information**
Pure premium for a class	←→	Pure Premium for all classes
Loss ratio for an individual risk	←→	Loss ratio for entire class
Indicated rate change for a territory	←→	Indicated rate change for entire state
Indicated rate change for entire state	←→	Trend in loss ratio

Suppose you are trying to estimate the indicated rate change for a territory within a state, but your company has a limited volume of business in the territory. An option may be to weight the indicated change from territorial data alone with the indicated change for the entire state. This way you have reflected territorial experience in your rate change to the extent that it is credible.

The loss ratios shown below in Table 3 were produced in a computer simulation that modeled the insurance random loss process. The expected loss ratio is 60 for both the small and big states, but the observed (simulated) loss ratios will randomly vary around this value. As we would expect, the variation is much larger for the small state. In the larger state the loss ratio hovers around 60 in each year. Five-year average loss ratios were calculated and then state indicated rate changes were calculated using the expected loss ratio of 60 as the permissible loss ratio. For example, in the small state -28.3% = (43/60 - 1.000). Using one of the formulas that we will discuss in a moment, credibility values Z were calculated for each state.

TABLE 3				
	Small State		**Large State**	
	Earned ($000)	**Loss Ratio**	**Earned ($000)**	**Loss Ratio**
1990	69	17	7,100	58
1991	71	109	7,120	58
1992	72	62	7,180	60
1993	74	7	7,200	58
1994	74	19	7,400	61
Total	360	43	36,000	59
Permissible Loss Ratio		60		60
State Indication		-28.3%		-1.7
Credibility		10%		100%

Perhaps this data comes from a line of insurance that has an aggressive insurance to value program such that the inflationary trend in losses is exactly offset by the annual increases in the amount of insurance. In this case the trend in our loss ratio would be 0%. (For our data, we know that the trend in the loss ratio is 0% because each year has an expected loss ratio of 60.) We will apply our complement of credibility factor (1-Z) to this information. So, we would get the following two indications:

small state: .10 X [-28.3%] + (1 - .10) X [0.0%] = -2.8%

large state: 1.00 X [-1.7%] + (1 - 1.00) X [0.0%] = -1.7%

In both cases we know the right answer! We should take a 0.0% rate change in each state because our expected loss ratios are what we used for the permissible loss ratios. But, *because of the randomness inherent in our data, our indications are slightly off the mark.*

The important thing in the prior example is that we greatly improved the accuracy of our rate indication in the small state by incorporating credibility. We gave only a 10% weight to the raw indication arising from the small state's loss ratio. This had the result of dampening the effect of the randomness. To the extent possible we would like to use our observed data to calculate our estimate rather than rely on supplementary data, but given the randomness present in our observations, we need to temper the data. Using credibility theory we weight an estimate based on limited data with data from other sources. We want to find a weight Z that allows us to rely on our limited data to the extent reasonable, but which also recognizes that our limited data is variable.

There are two widely used formulas for the credibility Z as shown side by side in Table 4. For the classical credibility formula, if n > N then Z is set equal to 1.00. In the case of Bühlmann credibility, Z asymptotically approaches 1.00 as n goes to infinity.

TABLE 4	
Classical Credibility	Bühlmann credibility
$Z = \sqrt{\frac{n}{N}}$	$Z = \frac{n}{n+K}$
Also called:	Also Called:
(1) Limited Fluctuation Credibility	(1) Least Squares Credibility (2) Empirical Bayesian Credibility (3) Bayesian Credibility

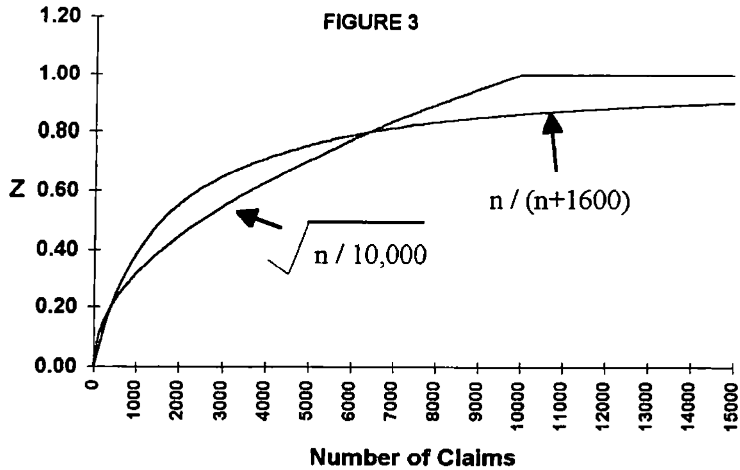
In both formulas n is a measure of the size of the body of data and is an indicator of the variability of the loss ratio or pure premium calculated from the data. n can be any of the following:

- number of claims
- amount of incurred losses
- number of policies
- earned premium
- number of insured unit-years.

These are not the only possibilities for n, but n needs to be some measure that grows directly with the size of the body of data that we have collected.

In practice both of the formulas can give about the same answer if N and K are chosen appropriately as displayed in Figure 3. Note that in the classical credibility case, when n is greater than or equal to 10,000, Z is identically 1.00.

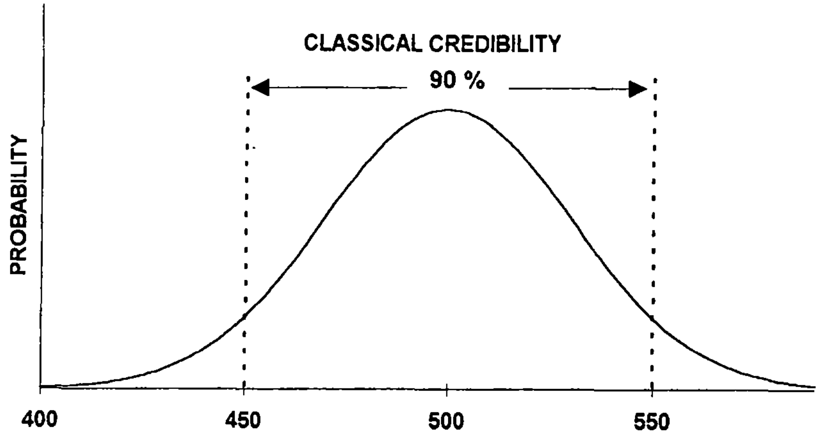
FIGURE 3



Classical Credibility

First we will discuss the classical credibility formula. Classical credibility attempts to restrict the fluctuation in the estimate to a certain range. N is calculated such that for fully credible data with n=N and Z=1.00, the observed pure premium or loss ratio will fall within a band about the expected value a specified percentage of the time. This is illustrated in Figure 4.

FIGURE 4



If N=5,200 claims, then the observed Pure Premium is within 10% of the "true" value 90% of the time.

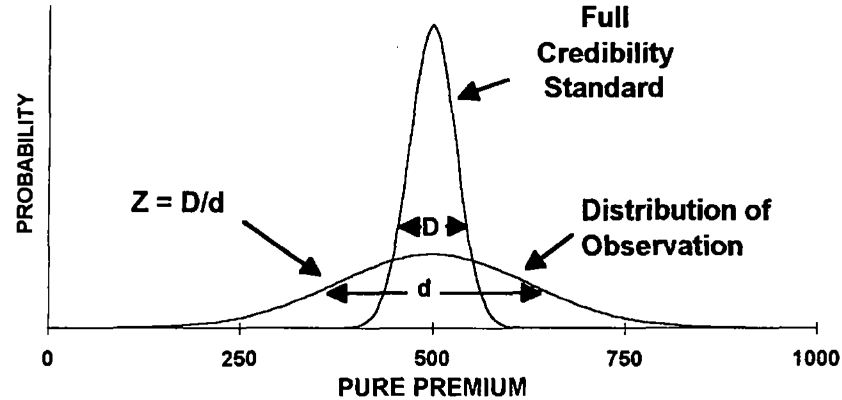
In this example the measure of the size of the body of data is the expected number of claims. When our body of data is large enough so that we expect 5,200 claims in our observation period, the observed pure premium will fall within k=10% of the true value P=90% of the time; that is, 90% of the time our pure premium calculated from our body of data will fall into the interval [450,550]. Both the 90% probability and the 10% width of the range must be selected by the ratemaker. If you wanted much less variance in your estimate you might select a P=99% probability and a k=2.5% error in your estimate. Of course, it would require a much larger body of data in the observation period to achieve this level of certainty.

The full credibility standard N is a function of the selected P and k values. A larger P value results in a larger N and a smaller k also produces a larger N. In order to calculate the N that corresponds to the selected P and k, one needs to make certain assumptions and also know something about the loss process. In classical credibility one assumes that the frequency of claims can be modeled by a Poisson distribution. Also, one needs an estimate of the average claim size and the variance in claim sizes. Using these an estimate of the variance in total losses can be computed. The next assumption is that the distribution of the total losses is normal, i.e. bell-shaped. Then, the N value can be calculated. This is all covered in much detail in the syllabus material for the actuarial exam that tests credibility theory.

One does not have to use the number of claims in the classical credibility formula, but instead can use earned premium, number of policies, or some other basis. We could convert our formula developed above to an earned premium basis. Suppose that in reviewing our data we calculate that on average there is approximately $2,500 in earned premium for each claim; that is, the ratio of earned premium to the number of claims is $2,500. A full credibility standard of (2,500 dollars/claim) x (5,200 claims) = $13,000,000 could be used in place of the 5,200 claims. Then, the credibility assigned to any data could be calculated from the earned premium of the data.

To calculate the full credibility standard, the denominator in the formula, the amount of variability acceptable in fully credible data must be defined by the selection of P and k values. For less than fully credible data the square-root formula determines the credibility Z. Figure 5 displays graphically the calculation of partial credibility.

FIGURE 5



In the graph the width of the curve representing the variability of data which just meets the standard for full credibility is represented by D. D can be considered the standard deviation of the curve. (If you prefer, D can be two standard deviations.) Likewise, d is the width corresponding to a smaller body of data that is less credible. It turns out that the credibility that should be assigned to the smaller body of data in this model is $Z = D/d$, the ratio of the standard deviation of the pure premium of the fully credible data to the standard deviation of the pure premium of the partially credible data. We will allow a standard deviation of size D, but if our body of data has a standard deviation of d, then we apply a weight of D/d to the data. If the pure premium (p.p.) calculated from the data is expected to have a standard deviation of d, then the quantity Z x (p.p) has a standard deviation of D, which is our target.

Bühlmann Credibility

The least-squares credibility model uses the credibility formula:

$$Z = n/(n + K)$$

K is defined by the following intimidating expression:

$$K = \frac{\textit{Expected Value of the Process Variance}}{\textit{Variance of the Hypothetical Means}}$$

A good way to think about least-squares credibility is in the context of experience rating where the rate charged to an insured is a manual rate modified to reflect the experience of the individual insured. The losses incurred by an insured are random, so an insured's loss ratio will fluctuate. The term "process variance" is the variance in the loss ratio of the risk. The "expected value of the process variance" is the average value of the variance across the risks within the population. Since each risk is unique, the expected loss ratios of the individual risks at the manual rates will vary across the population because the manual rates are based on averages calculated for groups of risks who are classified alike in the rating plan. Each risk has it's own "hypothetical mean" loss ratio. The "variance of the hypothetical means" is the variance across the population of risks of their individual hypothetical mean loss ratios.

In Figure 6 there are two risks, risk #1 and risk #2, each with its own loss ratio distribution curve. The process variance is a function of the width of the curve indicated by the [1] in the figure. As mentioned above the width of the curve can be thought of as some multiple of the standard deviation. The process variance is the square of the standard deviation. So the wider the curve, the larger the process variance. [2] marks the difference in the hypothetical means between the risks. The variance in the hypothetical means across the population is a function of the differences in the hypothetical means between the risks.

When the process variance of the risks is large in relation to the difference in the means of the risks, K is large. A large K means that the credibility $Z = n/(n + K)$ is small. Looking at the second graph in Figure 6, we see that there is a broad band where the two risks' loss ratios overlap. Since the loss ratio of each risk is so variable, it makes sense to give more weight to the manual rate calculated from the average experience of a large group of similar risks and less weight to the experience of the individual risk.

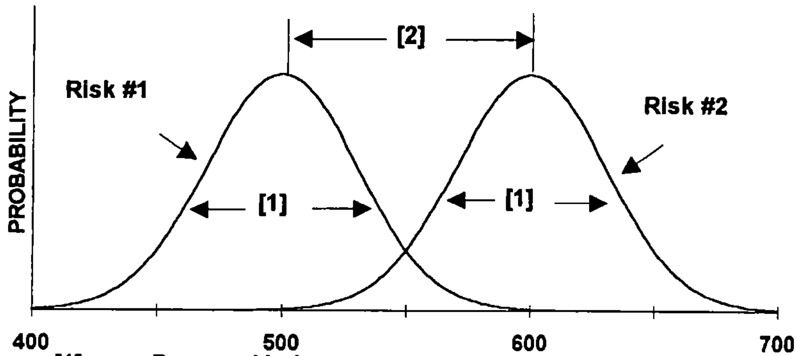
Small process variances in relation to the differences in the means of the risks results in a small K value and a larger credibility Z. This scenario is represented by the bottom graph in Figure 6. The distributions of the two risks do not overlap. The larger credibility Z means more weight is assigned to the experience of the individual risk and less, (1-Z), to the experience of the population.

Several Examples

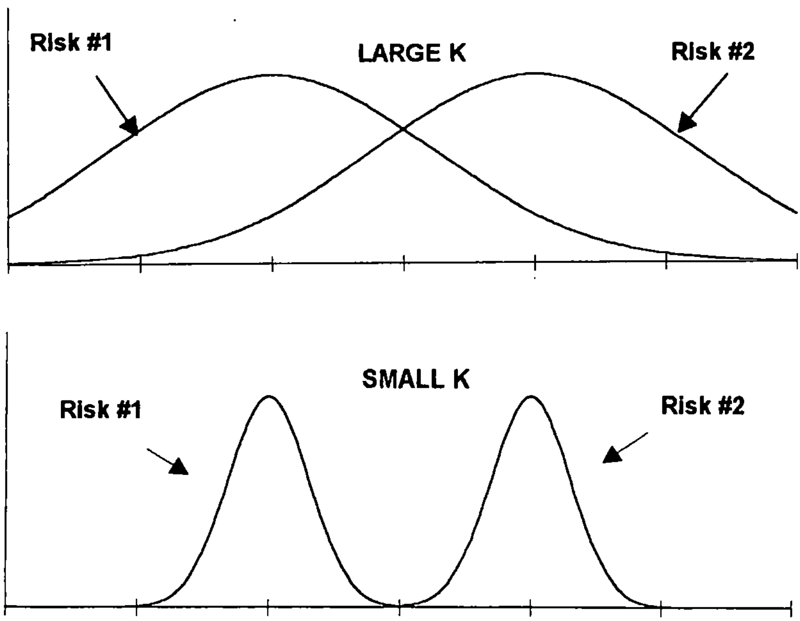
Examples of credibility formulas developed by the Insurance Services Office are displayed in Table 5. The first set of formulas are used in Homeowners ratemaking and are based on the classical credibility model. The measure of the size of the body of data and its consequent variability is in the units of house-years; that is, one house insured for one year contributes one unit. In making a statewide change 240,000 house-years are required for full credibility, and with that large of a body of data, the observed experience should be within 5% of the actual value 90% of the time. In computing territorial changes within the state, 60,000 house-years are assigned full

FIGURE 6

CALCULATION OF K



[1] --- Process Variance
[2] --- Variance of Hypothetical Means



credibility and the observed territorial experience is expected to be within 10% of the expected value of 90% of the time. As stated previously, the actuary needs to decide on the units for n, the size of the P value, and the size of the k value.

TABLE 5	
Credibility Formulas **Insurance Services Office**	
Homeowners: Owners Forms	
Statewide Changes $Z = \sqrt{\frac{n}{240,000}}$ n = house years 90% confident within 5% of actual value	*Territorial Changes* $Z = \sqrt{\frac{n}{60,000}}$ n = house years 90% confident within 10% of actual values
Manufacturers & Contractors	
Statewide Changes $Z = \sqrt{\frac{n}{8,000}}$ n = number of occurrences in three years 90% confident within 7% of actual value	*Relativities* $Z = \sqrt{\frac{n}{25,000}}$ n = number of occurrences in five years 95% confident within 5% of actual values
General Liability Experience Rating	
$Z = \frac{L}{L + \$177,000}$ L = expected loss costs (including ALAE) at $100,000 basic limits	

The next set of formulas in Table 5 are used by ISO in Manufacturers & Contractors ratemaking. Statewide changes require 8,000 claims (occurrences) in a three-year period, and with this many expected claims, the experience of the body of data should be within 7% of the expected value 90% of the time. The full credibility standard for relativities within M&C, such as class relativities, is much tougher with 25,000 claims required for a P=95% and k=5%.

The selection of P and k is probably more art than science. If the body of data that the actuary is working with is of limited size and there is no good surrogate for the data to which to assign the complement of credibility, then the actuary may select a smaller P and larger k to produce a smaller requirement for full credibility. If the actuary wants to make the rates more responsive to current experience he or she may also select a smaller P and a larger k. If rate stability is the most important goal then larger P and smaller k may be selected.

The last formula in Table 5 is the credibility to be assigned to an individual insured's data in General Liability experience rating and it is based on the Bühlmann model. In a loss cost environment, L represents the expected loss costs (expected incurred losses and allocated loss adjustment expenses) for the individual risk. Before the advent of loss costs, premium designated by E was used instead of L. The expected loss costs included in L are $100,000 basic limits losses. ISO has recently converted from $25,000 basic limits to $100,000. At $100,000 basic limits it was necessary to increase the K value in the denominator to $177,000 from its previously smaller value that applied when $25,000 basic limits losses were used in computing the experience rating adjustment. If unlimited losses were used in the experience rating formula, then an even larger K value would be necessary because the expected value of the process variance would become even larger.

Reducing Variability of the Data

The data used by ratemakers in the insurance business arises from a random process; in fact, it is this randomness that makes insurance necessary. The ratemaker is confronted with the task of finding the proper premiums to charge insureds without knowing for sure what the cost will be to the company to provide the insurance. The ratemaker estimates the cost of future payments in insurance claims by his or her company by analyzing past costs. The ratemaker wants to use the most relevant data to estimate future costs, but he or she must also deal with the variability inherent in the data.

One way to decrease the variability in ratemaking data is to use a larger body of data. Here are several ways to do this:

- include more years in the experience period
- use Bureau data
- combine data into fewer, but larger groups

Each of these involves a tradeoff. If more years are included in the experience period then it becomes necessary to apply larger trend factors to the older data and trend can be tough to estimate. Also, the book of business to which new rates will apply may be different from the business that produced the experience years ago. The same goes for Bureau data. The insureds included in Bureau data may be very different from the average insured in the ratemaker's data. Combining the data into fewer, but larger groups, may limit a company's ability to effectively compete against competitors who can better identify the proper price to charge an insured.

Another approach to decreasing the variability in losses used in ratemaking is to:

- cap large losses
- remove catastrophes

Of course, if we do either of the above we must put something back to make up for the losses we removed. One method to cap large losses is to do basic limits ratemaking by state, territory, class, etc., and calculate basic limits rates. Then, rates for higher limits are computed using increased limits factors calculated based on the aggregate data for many states and classes. Another approach is to limit all losses at some set amount, for example $150,000, and then to prorate the excess losses amount back by state, territory, class, etc. Catastrophe losses can be removed from the data and a catastrophe load substituted in its place. This load can be computed from a very long observation period, thirty years or more for weather losses, or a computer model that attempts to model the catastrophe loss process.