A GRAPHICAL ILLUSTRATION OF EXPERIENCE RATING CREDIBILITIES

HOWARD C. MAHLER

Abstract

This paper combines a simple experience rating example with a set of graphs in order to illustrate key credibility concepts as they relate to experience rating. As part of this graphical approach, credibility will be related to linear regression.

ACKNOWLEDGEMENT

I wish to thank Silvia Fitcher and Dotty Culleton for typing this paper.

1. INTRODUCTION

Philbrick [1] uses his excellent target shooting example to graphically illustrate some key concepts of credibility. Hewitt [2] uses a die/spinner example to illustrate important ideas of credibility. In this same spirit, this paper will combine a simple experience rating example with a set of graphs to illustrate key credibility ideas as they relate to experience rating. As part of the graphical approach, credibility ideas will be related to linear regression.

Prior and subsequent experience will be simulated for various sets of insureds for different sets of simple assumptions. This simulated data for the various examples will be used to illustrate that the slope of the regression line between prior and subsequent experience is one estimate of the Bühlmann credibility. Finally, these same examples will be used to illustrate that the expected squared error between the actual and predicted subsequent experience is minimized when the weight given to the observed experience is equal to the Bühlmann credibility.

2. EXPERIENCE RATING

The goal of experience rating is to use an individual insured's experience to help predict future loss costs.[1] If the individual risk's experience were observed to be worse than average, we would predict that his future experience would also likely be somewhat worse than average. Therefore, we would be likely to charge this insured somewhat more than average.

Credibility, as used in experience rating, quantifies how much worse or better an insured's future experience is expected to be based on a particular deviation from average observed in the past. In the simplest case:[2]

$$\begin{aligned}\text{New Estimate} &= (\text{Credibility})(\text{Observation})\\ &\quad + (1 - \text{Credibility})(\text{Overall Mean})\\ &= (\text{Overall Mean}) + (\text{Credibility})\\ &\quad \times (\text{Observation} - \text{Overall Mean}).\end{aligned}$$

In Appendix A, Bühlmann credibilities, Z, are calculated for various situations, using the formulas:

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

$$K = EPV/VHM$$

[1] See, for example, Meyers [3], Mahler [4], Finger [5], Gillam and Snader [6], and Tiller [7].

[2] The actual applications have a number of complications beyond the scope of this paper.

where

Z = Bühlmann credibility,

N = Number of years of data (from a single insured),

K = Bühlmann credibility parameter,

EPV = Expected value of process variance for a single unit of the risk process (i.e., for one insured for one year), and

VHM = Variance of the hypothetical means for a single unit of the risk process (i.e., for one insured for one year).

3. SIMPLE EXAMPLE

The following very simplified assumptions will be used in various combinations to illustrate credibility ideas. See Table 1 for a summary of the different situations illustrated.

TABLE 1

SUMMARY OF DIFFERENT SITUATIONS

Situation Number*	Quantity of Interest	Types of Insureds	Figure Number(s)	Credibility Estimated	Credibility Theoretical
1	Frequency	50 Good, 50 Bad	1, 2	40%	33%
1	Frequency	3 Years of Prior Data 50 Good, 50 Bad	3	58%	50%
2	Frequency	50 Excellent, 50 Ugly	4, 5	78%	81.8%
3	Frequency	50 Excellent, 50 Good, 50 Bad, 50 Ugly	6, 7	72%	71.4%
4	Unlimited Losses	125 Excellent, 125 Ugly	8, 9	51.5%	52.9%
5	Limited Losses	125 Excellent 125 Ugly	10	71.4%	70.1%

*See Appendix A for more details.

Claim frequency for individual insureds is assumed to be Poisson.[3] Claim severity is assumed to be given by a Pareto distribution[4] with shape parameter 3 and scale parameter 20,000. Frequency and severity are assumed to be independent. There are four possible types of insureds with different Poisson parameters:

Type	Average Annual Claim Frequency
Excellent	5
Good	10
Bad	15
Ugly	20

In Appendix A, the usual Bühlmann credibility techniques have been applied to various situations involving these four types of insureds in order to quantify the credibility to be assigned to the past experience of an insured. A set of graphs has been constructed to illustrate these same situations.

These graphs illustrate the connection between Bühlmann credibility and least squares linear regression. For the simple situations dealt with here, the slope of the least squares regression line between the past and subsequent observations of insureds is an estimate of the Bühlmann credibility. Appendix B provides a mathematical demonstration of this relationship. Not only is this relationship approximate,[5] but the slope from the regression will vary in particular examples due to random fluctuations. Thus, the estimated credibility will not exactly equal the theoretical Bühlmann credibility.

4. GRAPHS OF FREQUENCY EXAMPLES

Assume we have 100 insureds all in the same risk classification, territory, etc. The first graph, Figure 1, shows simu-

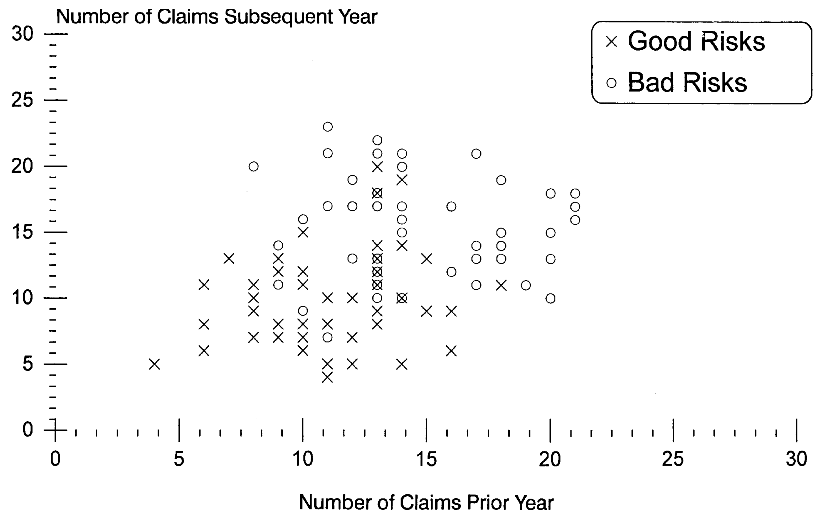
[3]The Poisson parameter for each insured stays the same over time.

[4]$F(x) = 1 - (20{,}000/(20{,}000 + x))^3$.

[5]As derived in Appendix B, one determines the expected value of a numerator and denominator separately and then assumes that $\mathrm{E}[A/B] \approx \mathrm{E}[A]/\mathrm{E}[B]$ in the sit-

FIGURE 1

SIMULATED CLAIMS EXPERIENCE



Situation 1: 50 Good Risks (Poisson 10) and 50 Bad Risks (Poisson 15)

lated claim counts for these 100 insureds divided into two equal groups. In this graph, the "Good Risks" are labeled with crosses and the "Bad Risks" with circles. In both the real world[6] and many of the subsequent graphs, the risks come without such labels attached. (If they did come with such labels, we would not need to use credibility.)

The 50 Bad Risks each have an expected claim frequency of 15 while the 50 Good Risks each have an expected claim frequency of 10. For each of the 100 insureds, a single prior year of simulated claim counts has been plotted against a single subsequent year of simulated claim counts. For example, one of

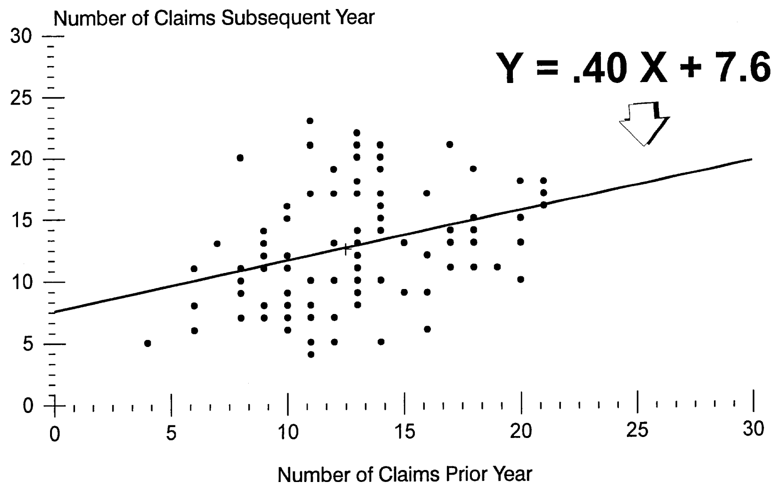
---

uations to which the result is being applied. In general, $E[A]/E[B]$ is not an unbiased estimator of A/B.

[6]In the real world, there is no way to precisely determine any individual's expected future frequency.

FIGURE 2

SIMULATED CLAIMS EXPERIENCE
GOOD AND BAD RISKS



Situation 1: 50 Good Risks (Poisson 10) and 50 Bad Risks (Poisson 15)

the Good Risks had 4 claims in the prior year and 5 claims in the subsequent year. This is indicated by a cross at the point (4,5). There is considerable overlap between the groups. Nevertheless, the Good Risks are more likely to be in the lower left while the Bad Risks are more likely to be in the upper right of the graph.

The next graph, Figure 2, shows the same 100 insureds without labels. In Figure 2 a least squares regression line has been fit to the points. One could use this fitted line to predict a future year's experience based on an observation. Since the line slopes upwards, a worse than average former year would lead one to predict a worse than average subsequent year.

So if one observed 20 claims in a year for an insured, one might predict about 15 claims for that insured next year, compared to the overall average of 12.5. The formula for this least squares line is approximately:

$$Y = .40X + 7.6.$$

The equation can be restated in the form of the "basic credibility formula:"

$$\text{Estimate} = Z(\text{Observation}) + (1 - Z)(\text{Overall Mean}),$$

with the credibility $Z = 40\%$ and

$$(1 - Z)(\text{Overall Mean}) = (60\%)(12.5) = 7.5 \approx 7.6.$$

With only 100 insureds, this result is subject to considerable random fluctuation. Thus, the estimated credibility of 40% is not equal to the theoretical Bühlmann credibility. The simulation with many more insureds would give a credibility of 1/3, the theoretical value as shown in Appendix A, Situation 1.

The credibility is just the slope of the straight line. It is the weight given to the observation.

Note the way that the fitted line passes through the point (12.5, 12.5), denoted by a plus. Average experience in the prior year yields an estimate of average experience in the subsequent year. This follows from rewriting the basic credibility formula as Estimate = Overall Mean + Z(Observation – Overall Mean).

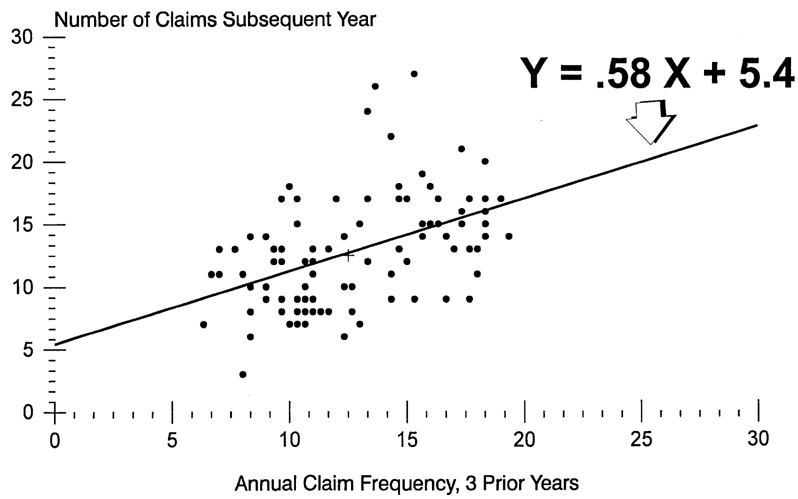
Note that the line $Y = X$, with a slope of unity, would correspond to 100% credibility, while the line $Y = 12.5$ with a slope of zero, would correspond to zero credibility. In general, the slope and the Bühlmann credibility will be between zero and one.

These general features displayed in Figure 2 will carry over to subsequent figures. The least squares line will slope upwards and pass through the point denoting average experience in the prior and subsequent period. The slope will be (approximately) equal to the credibility.

The next graph, Figure 3, is similar to Figure 2 but shows *three* years of prior experience rather than one. Note that the X-axis is

FIGURE 3

SIMULATED CLAIMS EXPERIENCE, 3 PRIOR YEARS
GOOD AND BAD RISKS



Situation 1: 50 Good Risks (Poisson 10) and 50 Bad Risks (Poisson 15)

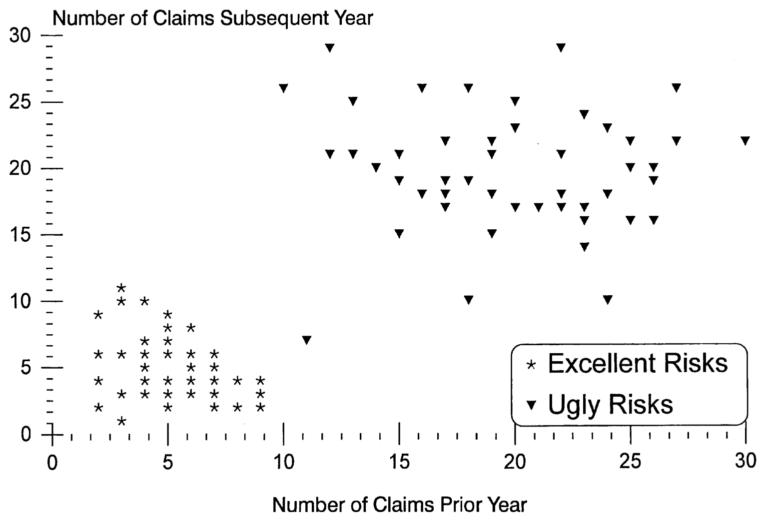
now the *annual* claim frequency observed over three years. We expect three years of data to contain more useful information and thus be given more weight than would one year. In fact, a fitted straight line has a larger slope of about 60% (actually 58%) corresponding to a credibility of 60%. One way to increase the credibility of data is to increase the volume of data.

In the case of Figures 2 and 3, the credibility is equal to $N/(N+K)$ where N is the number of years of data and $K = 2$. (See Appendix A, Situation 1.) This formula is used quite often, with the "Bühlmann credibility constant" K dependent on the statistical properties of the particular situation. Note that for Figure 2 with one year of prior data, $Z = 1/(1+2) = 33\%$, while in Figure 3 with three years of prior data, $Z = 3/(3+2) = 60\%$.

The next graph, Figure 4, shows 100 risks divided this time into Excellent Risks and Ugly Risks. The Excellent Risks are

FIGURE 4

SIMULATED CLAIMS EXPERIENCE



Situation 2: 50 Excellent Risks (Poisson 5) and 50 Ugly Risks (Poisson 20)

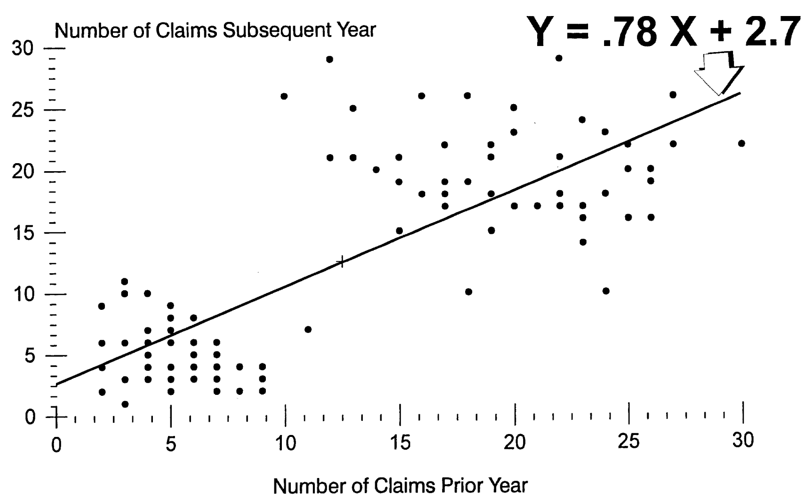
shown by asterisks and the Ugly Risks by wedges. The mean frequencies are 5 and 20 rather than 10 and 15 as in the previous exhibits. Therefore, the two groups are spread apart much more. Since there is more dispersion between risks,[7] each risk's data will be given more credibility than in the first graph.

This can be seen in the next graph, Figure 5, where a straight line has been fit to these points. The line has a much larger slope than the line in Figure 2, corresponding to higher credibility of about 82%. (The estimated credibility is 78%. Again the results of an experiment with only 100 risks differs from the theoreti-

[7]The experience is more likely to distinguish between excellent and ugly risks, than between good and bad risks. This is quantified via the variance of hypothetical means (*VHM*). As shown in Appendix A, the *VHM* in Situation 2 of 56.25 is much larger than that in Situation 1 of 6.25.

FIGURE 5

SIMULATED CLAIMS EXPERIENCE
EXCELLENT AND UGLY RISKS



Situation 2: 50 Excellent Risks (Poisson 5) and 50 Ugly Risks (Poisson 20)

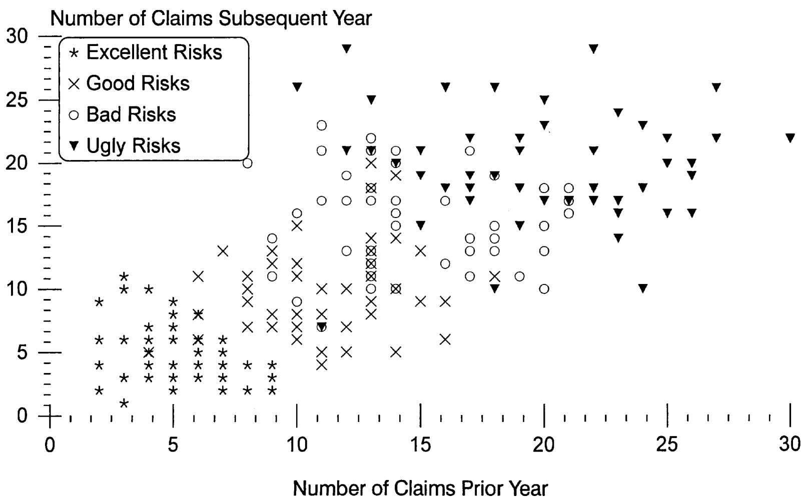
cal result of 81.8% in Appendix A, Situation 2, due to random fluctuation.) So due to the larger variation in hypothetical means (holding everything else equal) in Figure 5 versus Figure 2, the Bühlmann credibility increased from 33% to 82%. The value of the individual risk's information *increased relative* to the information contained in the overall mean. Conversely, the *relative* value of the information contained in the overall mean *decreased.*

The next graph, Figure 6, combines the four different types of insureds. This starts to approach the real world situations where risks' expected claim frequencies are assumed to be along a continuous spectrum, rather than being of discrete types.[8] We can see

[8]One could approach a continuous situation similar to the Gamma–Poisson frequency process. The Gamma–Poisson frequency process is explained, for example, in Hossack, Pollard and Zehnwirth [8], Herzog [9], or Mahler [10].

FIGURE 6

SIMULATED CLAIMS EXPERIENCE



Situation 3: 50 Excellent Risks (Poisson 5), 50 Good Risks (Poisson 10), 50 Bad Risks (Poisson 15), and 50 Ugly Risks (Poisson 20)

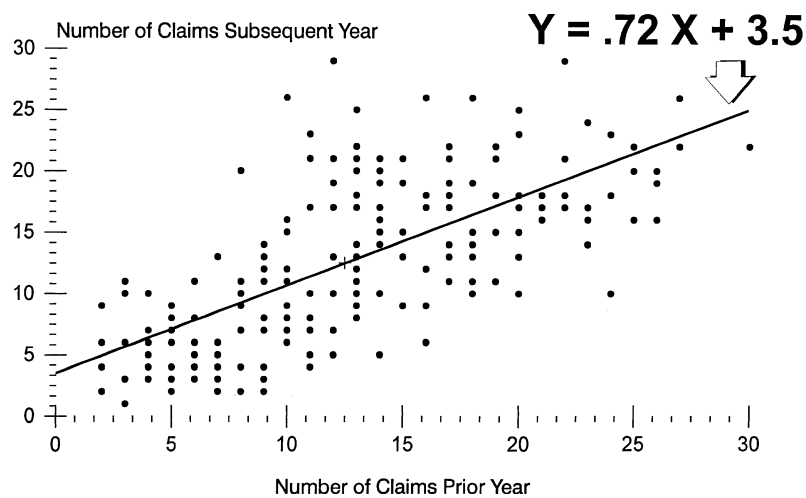
plenty of overlap between the four types of insureds, although since we labeled the insureds, we can discern the grouping of different types.

The next graph, Figure 7, shows a line fit to data from all four types. There the slope of 72% is between the slopes of 40% and 78% that we got when dealing with just two groups in Figures 2 and 5. All else being equal,[9] this makes sense since the variation of the hypothetical means is in between the variations of hypothetical means for those two situations. The theoretical credibility of 71% determined in Appendix A, Situation 3, is between the theoretical credibilities of 33% and 82% for Situations 1 and 2 which deal with only two groups.

[9] Specifically, the expected value of the process variance is the same in all three situations.

FIGURE 7

SIMULATED CLAIMS EXPERIENCE
EXCELLENT, GOOD, BAD, AND UGLY RISKS



Situation 3: 50 Excellent Risks (Poisson 5), 50 Good Risks (Poisson 10), 50 Bad Risks (Poisson 15), and 50 Ugly Risks (Poisson 20)

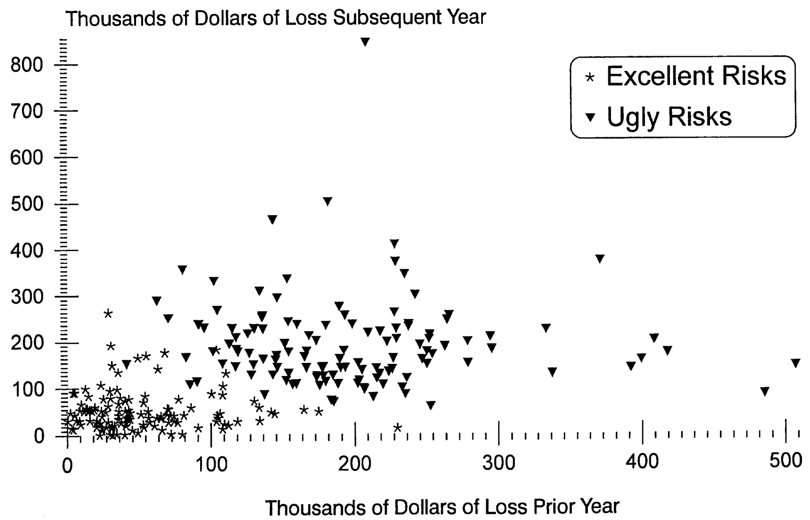
5. GRAPHS OF PURE PREMIUM EXAMPLES

The following graphs will all involve 125 Excellent and 125 Ugly Risks and not only deal with claim frequency, but with claim severity as well. By looking at dollars of loss rather than numbers of claims, as can be seen on the next graph, Figure 8, we introduce more random fluctuation.[10] Therefore, the relative value of the observation is less compared to the overall average; the credibility goes down. One way to *decrease* the credibility of data is to *increase* the variability of the data.

[10]In the absence of the labels, it would be somewhat easier to distinguish the Excellent and Ugly risks in Figure 4 dealing with frequency only than in Figure 8 dealing with dollars of loss.

FIGURE 8

SIMULATED LOSS EXPERIENCE



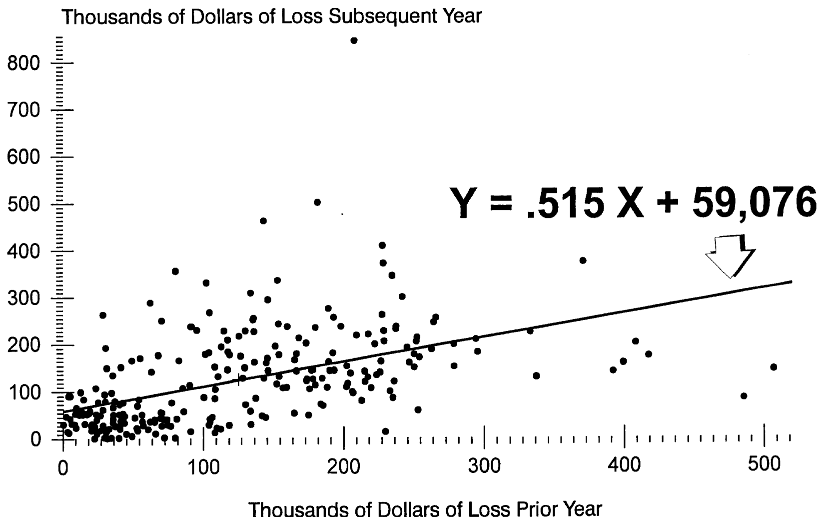
Situation 4: 125 Excellent Risks (Poisson 5) and 125 Ugly Risks (Poisson 20), Pareto Severity (3, $20,000)

As can be seen on the next graph, Figure 9, the slope of the fitted line is 51.5%. As shown in Appendix A, Situation 4, the theoretical credibility is 53% compared to 82% for the corresponding claim frequency Situation 2. The greater random fluctuation, which is quantified by the larger "process variance," has decreased the credibility assigned to the observations.

In practical applications, one often limits the size of claims entering into experience rating, since one way to decrease the variability of the data is to cap losses. The final graph in this series, Figure 10, shows the results of limiting each claim to $25,000. (This capping can be just for the purposes of experience rating or could involve an actual policy limit.) The slope of the fitted line between prior limited losses and subsequent limited losses is 71.4%. As determined in Appendix A, Situation 5,

FIGURE 9

SIMULATED LOSS EXPERIENCE



Situation 4: 125 Excellent Risks (Poisson 5) and 125 Ugly Risks (Poisson 20), Pareto Severity (3, $20,000)

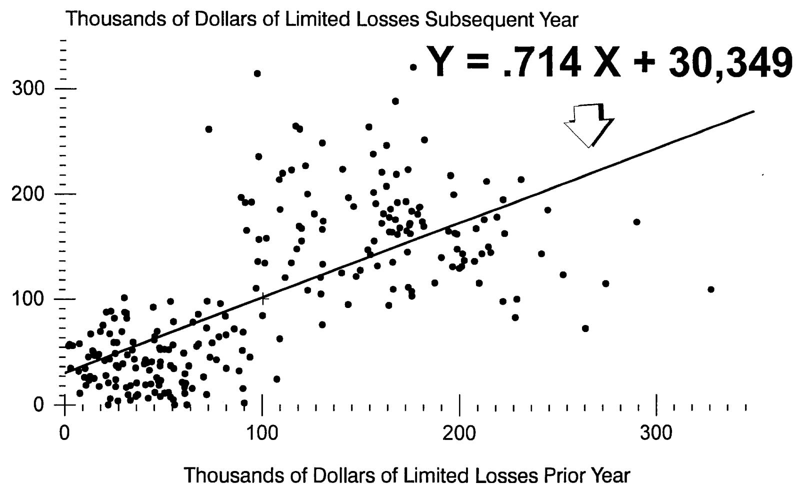
the theoretical credibility of 70% when using limited losses compares to 53% for total losses in Situation 4. Capping the losses has reduced the random fluctuations (i.e., has reduced the process variance) thereby increasing the credibility assigned to the experience. (Basic limit losses are less volatile than total limits losses.) For more on how to analyze experience rating plans, see for example Meyers [3] or Mahler [4].

6. EFFECT OF RANDOM FLUCTUATIONS ON ESTIMATED CREDIBILITIES

As mentioned above, the credibility estimated from regressing actual data sets will be affected by random fluctuations and, therefore, will not equal the theoretical Bühlmann credibility cal-

FIGURE 10

SIMULATED LOSS EXPERIENCE
Each Claim Limited to $25,000



Situation 5: 125 Excellent Risks (Poisson 5) and 125 Ugly Risks (Poisson 20), Pareto Severity (3, $20,000)

culated in Appendix A. The fewer insureds in the data set and/or the larger the process variance,[11] the larger is the impact from random fluctuations.

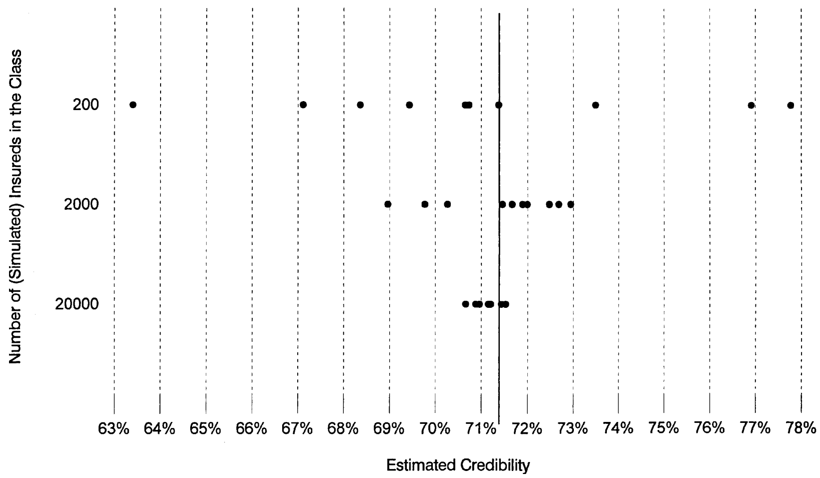
Figures 11 and 12 show the results of simulation experiments. Figure 11 deals with the frequency example with all four types of insureds as illustrated in Figures 6 and 7. The situation in Figure 7 with 200 insureds was simulated 10 separate times. This resulted in 10 different estimates of the credibility, ranging from 63.4% to 77.8%, as shown in Figure 11. Similar simulation ex-

---

[11] If the expected claim frequencies had been smaller, then the process variance would have been larger. For example, if the expected claim frequency for excellent risks were .05 rather than 5, one would need many more insureds to get as good an estimate of the credibility.

FIGURE 11

SIMULATION EXPERIMENTS
CREDIBILITIES ESTIMATED BY REGRESSION
CLAIM COUNTS
EXCELLENT, GOOD, BAD, AND UGLY RISKS



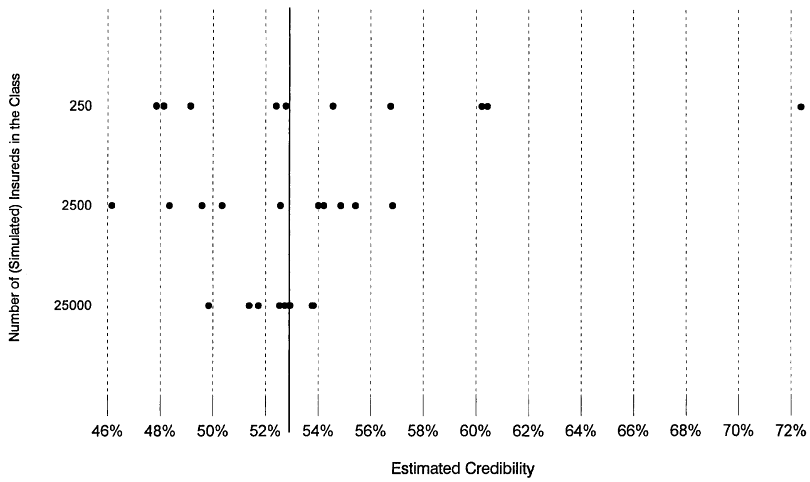
Credibilities are those to be applied to one observation of one insured. The theoretically correct value is 71.4%. Credibilities are estimated from the slope of the regression between one year of observations for the class and a subsequent year of observations for the class.

periments were performed for data sets of 2,000 and 20,000. As shown in Figure 11, with more insureds the credibility estimates are more tightly bunched and closer to the theoretically correct value.

Figure 12 is similar to Figure 11 but deals with the pure premiums rather than frequencies. With only 250 insureds there is considerable random fluctuation in the estimates. With 25,000 insureds the estimates are clustered between 50% and 54%. Due to the larger process variance, the estimates are less tightly clustered than they are in the examples involving frequency shown in Figure 11.

FIGURE 12

SIMULATION EXPERIMENTS
CREDIBILITIES ESTIMATED BY REGRESSION
LOSSES FOR EXCELLENT AND UGLY RISKS



Credibilities are those to be applied to one observation of one insured. The theoretically correct value is 52.9%. Credibilities are estimated from the slope of the regression between one year of observations for the class and a subsequent year of observations.

7. SQUARED ERRORS

Figures 13 through 17 illustrate the expected squared errors between the prediction and future observation for various weights applied to the observed data.

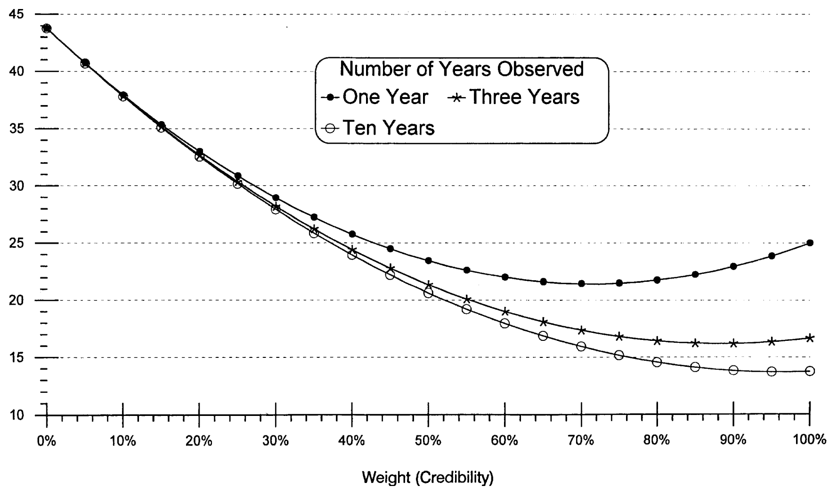
Figures 13 and 15 deal with the frequency example with all four types of insureds as illustrated in Figures 6 and 7. Figure 13 displays the expected squared error[12] as a function of the weight (credibility) given to the observed frequency. The expected squared error is a parabola as a function of the weight.[13]

[12]The expected value of the squared difference between the future observation and the prediction.

[13]This mathematical fact is demonstrated in Appendix C.

FIGURE 13

EXPECTED SQUARED PREDICTION ERRORS VS. WEIGHT GIVEN TO OBSERVED FREQUENCY
EXCELLENT, GOOD, BAD, AND UGLY RISKS



Equal numbers of: Excellent Risks (Poisson 5), Good Risks (Poisson 10), Bad Risks (Poisson 15), and Ugly Risks (Poisson 20). Expected value of process variance = 12.5, variance of the hypothetical means = 31.25. $K = 12.5/31.25 = 0.4$. Least squares credibilities are 71.4%, 88.2%, and 96.2%, for 1, 3, and 10 years of data, respectively.

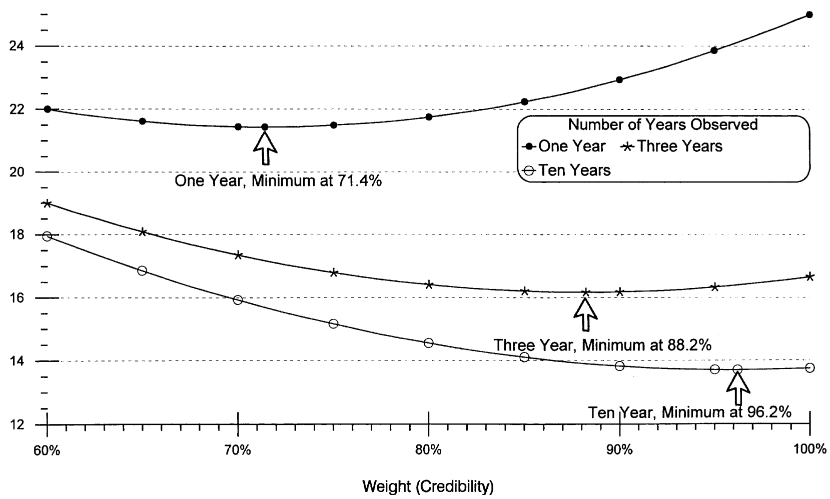
For one year of observed data, the expected squared error is minimized for a weight of 71.4%, the Bühlmann credibility for this situation. For three years of observed data, the minimum occurs for a weight of 88.2%. For ten years of observed data, the minimum occurs for a weight of 96.2%.[14]

As seen in Figure 13, as the number of years of observations increases, the prediction error from relying solely on the data (weight = 100%) declines, while the prediction error from relying solely on the a priori mean (weight = 0) remains the same. Thus, the place where the parabola reaches its minimum moves

[14]Note $10/(10 + .4) = 96.2\%$. Similarly $3/(3 + .4) = 88.2\%$ and $1/(1 + .4) = 71.4\%$.

FIGURE 14

EXPECTED SQUARED PREDICTION ERRORS VS. WEIGHT GIVEN TO OBSERVED FREQUENCY
EXCELLENT, GOOD, BAD, AND UGLY RISKS



Equal numbers of: Excellent Risks (Poisson 5), Good Risks (Poisson 10), Bad Risks (Poisson 15), and Ugly Risks (Poisson 20). Expected value of process variance = 12.5, variance of the hypothetical means = 31.25. $K = 12.5/31.25 = 0.4$. Least squares credibilities are 71.4%, 88.2%, and 96.2%, for 1, 3, and 10 years of data, respectively.

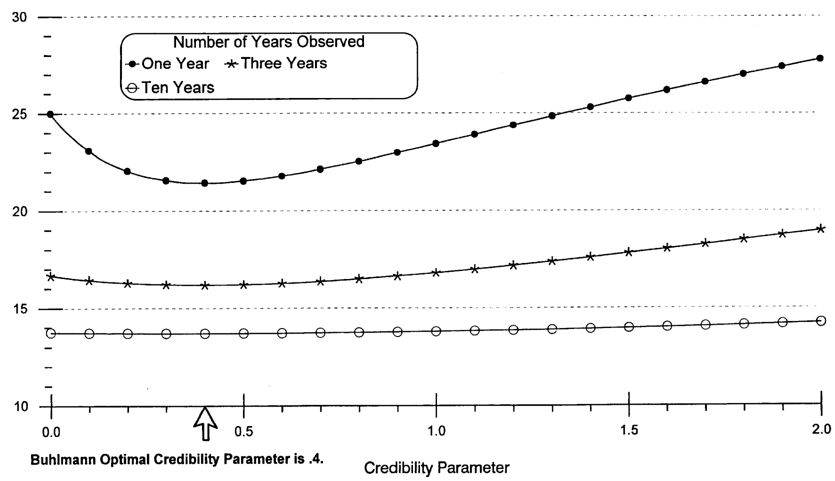
to the right as the number of years of data increases; the credibility increases becoming 100% in the limit as the number of years increases. For example, for one year of data the parabola reaches its minimum at 71.4%, while for three years of data the corresponding parabola reaches its minimum at 88.2%. Figure 14 is a magnified version of Figure 13, which more clearly displays the minima.

Figure 15 is similar to Figure 13, but here the expected squared error is displayed as a function of the "credibility parameter." In other words, we give N years of data weight $Z = N/(N+K)$, using the Bühlmann credibility formula with credi-

FIGURE 15

EXPECTED SQUARED PREDICTION ERRORS VS. CREDIBILITY PARAMETER USED TO DETERMINE WEIGHT GIVEN TO OBSERVED FREQUENCY

EXCELLENT, GOOD, BAD, AND UGLY RISKS



Equal numbers of: Excellent Risks (Poisson 5), Good Risks (Poisson 10), Bad Risks (Poisson 15), and Ugly Risks (Poisson 20). Expected value of process variance = 12.5, variance of the hypothetical means = 31.25, $K = 12.5/31.25 = 0.4$.

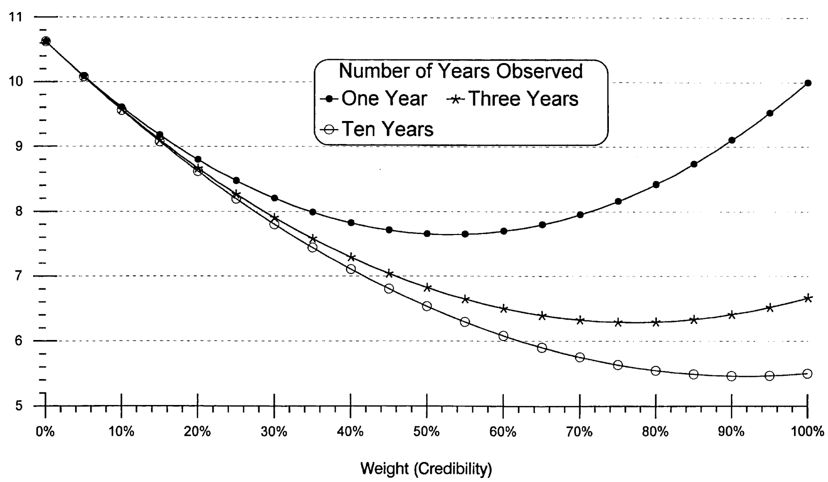
bility parameter K.[15] As shown in Appendix A, for Situation 3, the Bühlmann credibility parameter is 0.4; as seen in Figure 15, the expected squared error is indeed minimized for this value of the credibility parameter. Note the same credibility parameter of 0.4 is optimal regardless of the number of years of data observed.

Figures 16 and 17 are similar to Figures 13 and 15, but deal with the pure premiums rather than frequencies. Figure 16 shows the expected squared errors, which are parabolas as a function of

[15]In Figure 15 K is not necessarily the Bühlmann credibility parameter. Rather, we use a value of K to calculate a value of Z, which may not be the least squares Bühlmann credibility. In the case of Figure 15, 0.4 is the Bühlmann credibility parameter.

FIGURE 16

EXPECTED SQUARED PREDICTION ERRORS (BILLIONS) VS.
WEIGHT GIVEN TO OBSERVED FREQUENCY
EXCELLENT AND UGLY RISKS



Equal numbers of: Excellent Risks (Poisson 5) and Ugly Risks (Poisson 20). Expected Value of Process Variance = 5,000 million, Variance of the Hypothetical Means = 5,625 million. $K = 0.8889$. Least Squares Credibilities are 52.9%, 77.1%, and 91.8%, for 1, 3, and 10 years of data respectively.

the weight applied to the observed losses. Again, the expected squared errors are minimized when the weight given to the observed losses corresponds to the Bühlmann credibility.

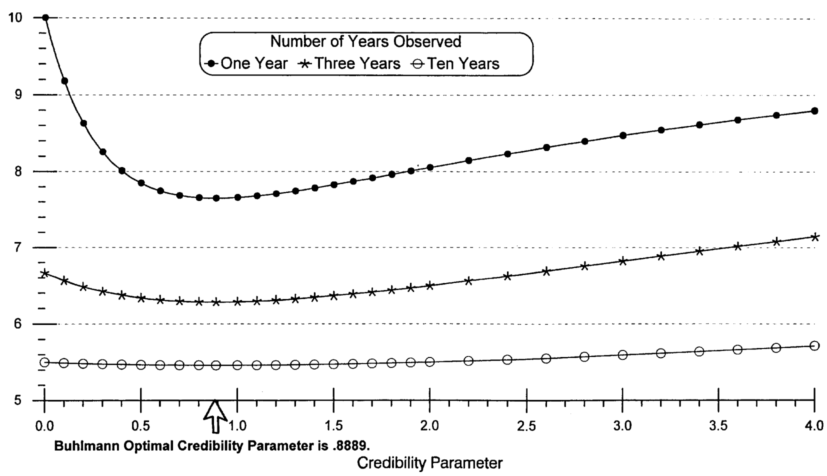
Figure 17 shows the expected squared error as a function of the credibility parameter. As shown in Appendix A, for Situation 4, the Bühlmann credibility parameter $K = .8889$. As seen in Figure 17, this value of the credibility parameter minimizes the expected squared errors.

8. CONCLUSIONS

Credibility, as used in experience rating, has been illustrated via graphs. The estimated credibility was equal to the slope of the line obtained from a least squares regression.

FIGURE 17

EXPECTED SQUARED PREDICTION ERRORS (BILLIONS) VS. CREDIBILITY PARAMETER USED TO DETERMINE WEIGHT GIVEN TO OBSERVED LOSSES EXCELLENT AND UGLY RISKS



Equal numbers of: Excellent Risks (Poisson 5) and Ugly Risks (Poisson 20). Expected value of process variance = 5,000 million, variance of the hypothetical means = 5,625 million. $K = 0.8889$.

Prior and subsequent experience has been simulated for various sets of insureds for different sets of simple assumptions. This simulated data for the various examples was used to illustrate that the slope of the regression line between prior and subsequent experience is one estimate of the Bühlmann credibility. Finally, these same examples were used to illustrate that the expected squared error between the actual and predicted subsequent experience is minimized when the weight given to the observed experience is equal to the Bühlmann credibility.

The regression technique shown here for illustrative purposes could be employed in simple situations. Where greater accuracy is desired or where the behavior is more complicated empirical Bayesian and other techniques have been developed to estimate credibilities from data.[16] In any case, the regression techniques applied to simulations of simple examples are another useful way to learn and understand the important basic ideas of credibility and experience rating.

[16]See for example ISO [11], Venter [12], Mahler [13], or Mahler [14].

REFERENCES

[1] Philbrick, Stephen W., "An Examination of Credibility Concepts," *PCAS* LXVIII, 1981, pp. 195–219.

[2] Hewitt, Charles C., "Credibility for Severity," *PCAS* LVII, 1970, pp. 148–171.

[3] Meyers, Glenn G., "An Analysis of Experience Rating," *PCAS* LXXII, 1985, pp. 278–317.

[4] Mahler, Howard C., Discussion of [3], *PCAS* LXXIV, 1987, pp. 119–189.

[5] Finger, Robert J., "Merit Rating for Doctor Professional Liability Insurance," *PCAS* LXXX, 1993, pp. 291–352.

[6] Gillam, William R. and Richard H. Snader, "Fundamentals of Individual Risk Rating," 1992, available from the CAS.

[7] Tiller, Margaret Wilkinson, "Individual Risk Rating," Chapter 3 of *Foundations of Casualty Actuarial Science*, Casualty Actuarial Society, 1990.

[8] Hossack, I. B., J. H. Pollard and B. Zehnwirth, *Introductory Statistics with Applications to General Insurance*, New York, Cambridge University Press, 1983.

[9] Herzog, Thomas N., *An Introduction to Credibility*, ACTEX, 1994.

[10] Mahler, Howard C., "A Student's Guide to Conjugate Priors," *New England Actuarial Seminars*, 1997.

[11] Insurance Services Office, "*Report of the Credibility Subcommittee: Development and Testing of Empirical Bayes Credibility Procedures for Classification Ratemaking*," September, 1980.

[12] Venter, Gary G., "Credibility," Chapter 7 of *Foundations of Casualty Actuarial Science*, Casualty Actuarial Society, 1990.

[13] Mahler, Howard C., "A Markov Chain Model of Shifting Risk Parameters," *PCAS* LXXXIV, 1997, pp. 581–659.

[14] Mahler, Howard C., "Credibility with Shifting Risk Parameters, Risk Heterogeneity and Parameter Uncertainty," *PCAS* LXXXV, 1998, pp. 455–653.

[15] Mayerson, Allen L., "A Bayesian View of Credibility," *PCAS* LI, 1964, pp. 85–104.

[16] Hewitt, Charles C., Review of [15], *PCAS* LII, 1965, pp. 121–127.

[17] Mahler, Howard C., "A Student's Guide to Bühlmann Credibility and Bayesian Analysis," *New England Actuarial Seminars*, 1997.

[18] Mahler, Howard C., "A Student's Guide to Classical Credibility," *New England Actuarial Seminars*, 1997.

[19] Mahler, Howard C., "A Student's Guide to Loss Distributions," *New England Actuarial Seminars*, 1997.

[20] Klugman, Stuart A., Harry H. Panjer and Gordon E. Willmot, *Loss Models: From Data to Decisions*, Wiley, 1998.

[21] Hogg, Robert V. and Klugman, Stuart A., *Loss Distributions*, Wiley, 1984.

APPENDIX A

CREDIBILITY FOR THE EXAMPLES

The formulas to be used are:[17]

$$K = EPV/VHM,$$

$$Z = N/(N+K),$$

where

K = Bühlmann credibility parameter,

EPV = Expected value of process variance for a single unit of the risk process (i.e., for one insured for one year,

VHM = Variance of the hypothetical means for a single unit of the risk process (i.e., for one insured for one year,

Z = Bühlmann credibility, and

N = Number of years of data (from a single insured).

The following information will be used in various combinations to illustrate credibility ideas.

Claim frequency for individual insureds is assumed to be Poisson.[18] Claim severity is assumed to be given by a Pareto distribution[19] with shape parameter 3 and scale parameter 20,000.Frequency and severity are independent. There are four possible types of insureds with different Poisson parameters:

Type	Average Annual Claim Frequency
Excellent	5
Good	10
Bad	15
Ugly	20

[17]These formulas are explained or derived in, for example, Mayerson [15], Hewitt [16], Hewitt [2], Philbrick [1], Herzog [9], Venter [12], and Mahler [17].

[18]The Poisson parameter for each insured stays the same over time.

[19]$F(x) = 1 - (20{,}000/(20{,}000 + x))^3$.

Situation 1: Frequency with Good and Bad Risks

A risk is selected at random from a class made up equally of Good and Bad Risks.

Since the frequencies are Poisson, their variance is equal to their mean. Therefore, the expected value of the process variance is equal to the overall mean = 12.5.

$$VHM = \{(10 - 12.5)^2 + (15 - 12.5)^2\}/2 = 6.25.$$
$$K = EPV/VHM = 12.5/6.25 = 2.$$

For one year of data (as in Figure 2), $Z = 1/(1+2) = 33\%$.

For three years of data (as in Figure 3), $Z = 3/(3+2) = 60\%$.

Situation 2: Frequency with Excellent and Ugly Risks

A risk is selected at random from a class made up equally of Excellent and Ugly Risks.

$$EPV = \text{overall mean} = 12.5.$$
$$VHM = \{(5 - 12.5)^2 + (20 - 12.5)^2\}/2 = 56.25.$$
$$K = EPV/VHM = .222.$$

For one year of data (as in Figure 5), $Z = 1/(1+.222) = 81.8\%$.

Situation 3: Frequency with Excellent, Good, Bad, and Ugly Risks

A risk is selected at random from a class made up equally of Excellent, Good, Bad, and Ugly Risks.

$$EPV = \text{overall mean} = 12.5.$$
$$VHM = \{(5 - 12.5)^2 + (10 - 12.5)^2 + (15 - 12.5)^2 + (20 - 12.5)^2\}/4 = 31.25.$$
$$K = 12.5/31.25 = .4.$$

For one year of data (as in Figure 7), $Z = 1/(1+.4) = 71.4\%$.

Situation 4: Pure Premiums (Unlimited Losses) for Excellent and Ugly Risks

A risk is selected at random from a class made up equally of Excellent and Ugly risks.

With a Poisson frequency,[20] the process variance of the pure premiums = (mean frequency)(second moment of the severity). (See, for example, Mahler [18].) Since the severity distribution is assumed to be the same for all risks, the expected value of the process variance = (overall mean frequency)(second moment of the severity).

For a Pareto distribution, $F(x) = 1 - (\lambda/(\lambda + x))^{\alpha}$, the second moment of the severity is $2\lambda^2/\{(\alpha - 1)(\alpha - 2)\}$, which in this case is 400 million. Therefore, since the mean frequency is 12.5, $EPV = (12.5)$ (400 million) = 5 billion.

For a Pareto distribution, the mean is $\lambda/(\alpha - 1) = 10{,}000$. Thus, the hypothetical mean pure premiums are 50,000 and 200,000. Thus, the $VHM = 5.625$ billion.

Thus, $K = EPV/VHM = 0.8889$.

For one year of data (as in Figure 9), $Z = 1/(1 + .8889) =$ 52.9%.

Situation 5: Limited Losses for Excellent and Ugly Risks

A risk is selected at random from a class made up equally of Excellent and Ugly Risks. One observes the losses limited to \$25,000 per claim and attempts to predict the future limited losses for the same insured.

[20] In general, for cases where frequency and severity are independent, the process variance of the pure premium = (mean frequency)(variance of severity) + (mean severity)^2 (variance of frequency). For a Poisson, mean frequency = variance of the frequency. Thus, the process variance of the pure premiums = (mean frequency)(variance of severity + mean severity^2) = (mean frequency)(second moment of severity).

For a Pareto distribution, $F(x) = 1 - (\lambda/(\lambda + x))^{\alpha}$, the limited second moment is given by[21]

$$\mathrm{E}[X^2;L] = \mathrm{E}[X^2]\{1 - (1 + L/\lambda)^{1-\alpha}[1 + (\alpha - 1)L/\lambda]\}.$$

In this case, $\mathrm{E}[X^2;25{,}000] = 400$ million $\{1 - (1 + 1.25)^{-2}[1 + (2)(1.25)]\} = 123.5$ million. As in Situation 4, *EPV* = (overall mean frequency)(second moment of the severity) = (12.5)(123.5 million) = 1.544 billion.

For a Pareto distribution, the limited expected value is given by[22]

$$\mathrm{E}[X;L] = \mathrm{E}[X]\{1 - (1 + L/\lambda)^{1-\alpha}\}.$$

In this case, $\mathrm{E}[X;25{,}000] = (10{,}000)(1 - (1 + 1.25)^{-2}) = 8{,}025$. Thus, the hypothetical mean pure premiums are (5)(8,025) = 40,125 and (20)(8,025) = 160,500. Therefore, *VHM* = 3.623 billion. K = *EPV/VHM* = 0.426. Note that while both the *EPV* and *VHM* declined compared to Situation 4, the *EPV* declined more. Therefore, the Bühlmann credibility parameter K declined from 0.8889 to 0.426. Thus, for one year of data (as in Figure 10) $Z = 1/(1 + .426) = 70.1\%$.

[21] See Mahler [19] or Klugman, Panjer, and Willmot [20].

[22] See Hogg and Klugman [21], Mahler [19], or Klugman, Panjer, and Willmot [20].

APPENDIX B

REGRESSION AND CREDIBILITY

It turns out that in the example here,[23] the Bühlmann credibility is approximately the slope of the least squares line between prior and subsequent observations, as will be shown in this Appendix. Also, it will be shown that the regression line is expected to pass approximately through the point (M,M), where M is the overall mean.

Let X_i be the prior observations (for one year) for the insureds in the portfolio and let Y_i be the subsequent observations (for one year) for the insureds in the portfolio. A regression line $y = ax + b$, with slope a and intercept b, can be fit between the prior and subsequent observations.[24] Then the slope of the least squares line is given by:[25]

$$a = \frac{(\Sigma X_i Y_i/O) - (\Sigma X_i/O)(\Sigma Y_i/O)}{(\Sigma X_i^2/O) - (\Sigma X_i/O)^2}.$$

Where O is the number of insureds observed.

The numerator has an expected value equal to the covariance of X_i and Y_i, the observations in two separate years.[26] This is assumed to be τ^2, the variance of the hypothetical means.[27]

The denominator has an expected value equal to the variance of the observations in a single year.[28] It is assumed that this

---

[23]This result holds in the case of the covariance structure assumed in Appendices A and C. In particular, there are no shifting risk parameters over time. More general covariance structures are discussed, for example, in Meyers [3], Mahler [4], Mahler [13], and Mahler [14].

[24]As is done in Figures 2, 3, 5, 7, 9, and 10.

[25]For simplicity we have assumed that each insured is of the same size and gets the same weight. Thus, we perform an unweighted regression.

[26]Recall that $\mathrm{Cov}[A,B] = \mathrm{E}[AB] - \mathrm{E}[A]\mathrm{E}[B]$.

[27]One of the assumptions underlying Bühlmann's credibility formula is that the covariance between different years of data is the variance of the hypothetical means.

[28]Recall that $\mathrm{Var}[A] = \mathrm{E}[A^2] - \mathrm{E}[A]^2$.

expected value is $\tau^2 + \eta^2$, the sum of the variance of the hypothetical means and the expected value of the process variance.[29]

If one plugs in the expected values of both the numerator and denominator, then we expect:[30]

$$a \approx \tau^2/(\tau^2 + \eta^2) = \text{Bühlmann credibility for one year.}^{31}$$

If X had been an observation for N years rather than one year, then the expected value of the process variance (of the frequency or pure premiums) would have declined by a factor of $1/N$; it would have been η^2/N rather than η^2 as for one year.[32] On the other hand, the variance of the hypothetical means would have remained the same.[33] Thus, with N years of data rather than one, the expected value of the numerator would have been the same, but the expected value of the denominator would have been $\tau^2 + \eta^2/N$ and

$$\begin{aligned} a &\approx \tau^2/(\tau^2 + \eta^2/N) = N/(N + \eta^2/\tau^2) = N/(N + K) \\ &= \text{Bühlmann credibility for } N \text{ years.} \end{aligned}$$

Thus, the slope of the regression line is approximately[34] equal to the Bühlmann credibility.

[29]The terms are each defined in terms of a single year of data. The total variance is equal to the *VHM* plus *EPV*.

[30]Note that this estimator which is a ratio of two unbiased estimators can be biased. This subject has been extensively discussed in relation to empirical Bayes credibility. See, for example, ISO [11] and Venter [12].

[31]$\tau^2/(\tau^2 + \eta^2) = VHM/(VHM + EPV) = 1/(1 + EPV/VHM) = N/(N + K)$, with $N = 1$ and $K = EPV/VHM$.

[32]See, for example, Mahler [17] or Mahler [14]. The process variance of the number of claims increases by a factor of N, since variances add for (independent) years. However, the claim frequency is the claim count divided by N, which introduces a factor of $1/N^2$ into the variances. The net result is a factor of $N/N^2 = 1/N$ for the process variance of the claim frequency.

[33]See, for example, Mahler [17] or Mahler [14]. The hypothetical annual means of the claim frequency are unchanged, thus, their variance is also unaffected. Alternately, the hypothetical mean claim counts are multiplied by N and, thus, their variance is multiplied by N^2. However, claim frequency is divided by N, which introduces a factor of $1/N^2$ into the variance. The *VHM* is, thus, multiplied by $N^2/N^2 = 1$.

[34]In general $E[A/B] \neq E[A]/E[B]$. Nevertheless, for the situations such as being dealt with here, $E[A/B] \approx E[A]/E[B]$.

Also, one can show that the regression line is expected to pass approximately through the point (M,M) where M is the overall mean. The intercept of the least squares line is

$$b = \frac{(\Sigma Y_i/O)(\Sigma X_i^2/O) - (\Sigma X_i Y_i/O)(\Sigma X_i/O)}{(\Sigma X_i^2/O) - (\Sigma X_i/O)^2},$$

where O is the number of insureds observed each year,

$\Sigma Y_i/O$ has an expected value equal to the overall mean M, and $\Sigma X_i^2/O$ has an expected value of the second moment of the average of N years of data. This is the sum of the variance of the average of N years of data plus the square of the overall mean. In turn, the variance of the average of N years of data[35] is equal to $\tau^2 + \eta^2/N$.

Thus, the expected value of $\Sigma X_i^2/O$ is equal to $\tau^2 + \eta^2/N + M^2$.

The numerator of a is equal to $\Sigma X_i Y_i/O - (\Sigma X_i/O)(Y_i/O)$. Thus, the expected value of $\Sigma X_i Y_i/O$ is equal to that of the numerator of a, τ^2, plus the expected value of $(\Sigma X_i/O)(\Sigma Y_i/O)$ which is (M) (M). Therefore, the expected value of $\Sigma X_i Y_i/O$ is $\tau^2 + M^2$.

$\Sigma X_i/O$ has an expected value equal to the overall mean M.

Thus, the expected value of the numerator of b with N years of data is

$$M(\tau^2 + \eta^2/N + M^2) - (\tau^2 + M^2)M = M\eta^2/N.$$

[35] Where $\mathbf{W}_j$ is the vector of data for year j

$$\begin{aligned}
\mathrm{Var}[(1/N)\Sigma \mathbf{W}_j, (1/N)\Sigma \mathbf{W}_k] &= (1/N^2)\mathrm{Var}[\Sigma \mathbf{W}_j, \Sigma \mathbf{W}_k] \\
&= (1/N^2)\Sigma\,\mathrm{Var}[\mathbf{W}_j, \mathbf{W}_j] + (1/N^2)\Sigma_{j\neq k}\mathrm{Var}[\mathbf{W}_j, \mathbf{W}_k] \\
&= (1/N^2)N(\tau^2 + \eta^2) + (1/N^2)((N^2 - N)(\tau^2)) \\
&= \tau^2 + \eta^2/N.
\end{aligned}$$

The denominator of b is the same as that of a, and has an expected value of $\tau^2 + \eta^2/N$ (for N years of data).

Thus, by substitution we expect

$$b \approx \frac{M\eta^2/N}{\tau^2 + \eta^2/N}.$$

Thus, since as was shown above,

$$a \approx \tau^2/(\tau^2 + \eta^2/N),$$

$$aM + b \approx M\tau^2/(\tau^2 + \eta^2/N) + (M\eta^2/N)/(\tau^2 + \eta^2/N) = M.$$

Thus, we indeed expect the regression line $y = ax + b$ to pass approximately through the point (M,M). Prior experience for an insured equal to the overall a priori expectation results in a prediction equal to the overall a priori expectation.

APPENDIX C

EXPECTED VALUE OF SQUARED ERRORS

This appendix discusses the expected value of the squared errors that result from the use of credibility to estimate an insured's future experience from the insured's past observed experience. The results of this appendix are illustrated in Figures 13 through 17. This appendix also shows how to calculate the Bühlmann credibility, which is the value that minimizes this expected squared error.[36]

Assume we have a time series, X_i, and we wish to estimate a future year of the same time series, $X_{N+\Delta}$, by weighting together observations X_i for $i = 1$ to N and the overall mean M. For example, the X_i could be the observed frequencies for a single insured over a series of individual years. If Z_i is the weight applied to year X_i, then

$$\text{Estimate} = \sum_{i=1}^{N} Z_i X_i + \left(1 - \sum_{i=1}^{N} Z_i\right) M.$$

Then the expected squared error comparing the estimate to the observation[37] is a quadratic function of the weights Z_i:[38]

$$V(Z) = \sum_{i=1}^{N}\sum_{j=1}^{N} Z_i Z_j C_{ij} - 2\sum_{i=1}^{N} Z_i C_{i,N+\Delta} + C_{N+\Delta,N+\Delta},$$

where $C_{ij} = \mathrm{Cov}[X_i, X_j]$.

[36]It turns out that this value of credibility also minimizes the squared error between the predictions and the true/hypothetical means and between the predictions and the Bayesian estimates. See for example Mahler [17].

[37]The expected squared error compared to the observation, with respect to the hypothetical mean, or with respect to the Bayesian estimate are each minimized by the value for credibility calculated using the formula derived in this appendix.

[38]See for example Mahler [13].

In the examples in Appendix A, the covariance structure is that underlying the Bühlmann credibility formulation:

$$C_{ij} = \tau^2 + \delta_{ij}\eta^2,$$

where C_{ij} = Covariance of year i and year j,

η^2 = Expected value of the process variance,

τ^2 = Variance of the hypothetical means, and

$\delta_{ij} = 1$ if $i = j$ and 0 if $i \neq j$.

Due to symmetry in this case, it turns out that the expected squared errors are minimized for $Z_i = Z_j$. Let $Z = \sum_{i=1}^{N} Z_i$ = total weight to be applied to N years of data. Then if $Z_i = Z_j = Z/N$, substituting into the formula for the expected squared errors:

$$\begin{aligned} V(Z) &= \sum_{i=1}^{N}\sum_{j=1}^{N} Z_i Z_j(\tau^2 + \delta_{ij}\eta^2) - 2\sum_{i=1}^{N} Z_i\tau^2 + \tau^2 + \eta^2 \\ &= Z^2\tau^2 + \eta^2(Z/N)^2 N - 2\tau^2 Z + \tau^2 + \eta^2 \\ &= Z^2(\tau^2 + \eta^2/N) - 2\tau^2 Z + \tau^2 + \eta^2. \end{aligned}$$

For Situation 3 in Appendix A, $\eta^2 = 12.5$ and $\tau^2 = 31.25$. Thus, for ten years of data $V(Z) = 32.5Z^2 - 62.5Z + 43.75$. This is one of the parabolas shown in Figure 13.

In order to minimize the expected squared error, we set the derivative $V'(Z) = 0$. This results in $Z = \tau^2/(\tau^2 + \eta^2/N)$ $= N/(N + \eta^2/\tau^2) = N/(N + K)$, where $K = \eta^2/\tau^2 = EPV/VHM$. For example, for 10 years of data in Figures 13 or 14, the parabola is minimized for $Z = 31.25/(31.25 + 12.5/10) = 0.962$. Alternatively, $K = 12.5/31.25 = 0.4$ and $Z = 10/(10 + .4) =$ 0.962. As seen in Figure 15, this value of K minimizes the expected squared error.