

Credibility Methods Applied to Life, Health, and Pensions





Credibility Applications for Life and Health Insurers and Pension Plans

Author

David B. Atkinson, FSA, MBA

President

DB Atkinson Consulting

TABLE OF CONTENTS

Preface	4
Purpose.....	4
Acknowledgements	4
Background and Scope	4
Chapter 1: Overview of Credibility	5
1.1 Credibility Methods	6
1.2 General Formula for Credibility Methods	7
Chapter 2: Limited Fluctuation Method	9
2.1 Relevance of Prior Rate	9
2.2 Square Root Formula.....	10
2.2.1 FULL CREDIBILITY.....	10
2.2.2 DERIVATION OF THE NUMBER 1,082	11
2.2.3 COUNTERPART OF 1,082 FOR MORTALITY AND LAPSE RISKS	12
2.3 Asymptotic Formula.....	13
2.4 Credibility Weights Based on Confidence Intervals	14
2.5 Mortality Study Example	16
2.5.1 MORTALITY STUDY DATA AND RESULTS	16
2.5.2 DETERMINATION OF THRESHOLD FOR FULL CREDIBILITY	17
2.5.3 CALCULATION OF CREDIBILITY-WEIGHTED RATE	18
2.6 Credibility Applied to Mortality Ratios	19
2.7 Strengths and Weaknesses	20
Chapter 3: Greatest Accuracy Method	21
3.1 Bayesian Credibility, Stepping Stone to Greatest Accuracy	21
3.2 Overview of Greatest Accuracy.....	21
3.3 Greatest Accuracy Formulas	22
3.4 Bühlmann-Straub Model	24
3.5 Estimating the Components of K	25
3.6 Strengths and Weaknesses	25
Chapter 4: Recommended Credibility Readings.....	26
4.1 Foundations of Casualty Actuarial Science.....	26
4.2 Credibility Practice Note, American Academy of Actuaries, July 2008.....	26
4.3 Credibility Theory Practices, SOA Report, December 2009.....	27
4.4 Other Recommended Readings for Life, Health, and Pensions	28
4.5 Other Recommended Readings from Casualty Actuarial Practice.....	30
Appendix A: Calculation of $E[Var[X]]$ and $Var[E[X]]$.....	34
A.1 Weights, Means, and Variance	34
A.2 Portfolio Statistics.....	36
Appendix B: Minimizing Mean-Squared Error	37
About the Society of Actuaries.....	40

Preface

Purpose

The purpose of this paper is to promote more informed applications of credibility methods for life insurers and pension plans. As an introduction to credibility, it allows a reader with a basic knowledge of probability and statistics to quickly grasp the usage and limitations of credibility methods. With this foundational understanding, readers should be able to apply basic credibility concepts and dig deeper into the literature on credibility. To that end, there are recommended readings in Chapter 4.

Acknowledgements

The Society of Actuaries would like to thank the following members of the project oversight group who provided valuable input and oversight for this paper. Special thanks are due to Stuart Klugman who devoted many hours to the early drafts, working closely with the author to enhance the accuracy of the paper.

- Mary Bahna-Nolan
- Cathy Bierschbach
- Paul Correia
- Steve Ekblad
- Chuck Fuhrer
- Tasha Khan
- Stuart Klugman
- Achilles Natsis
- Marianne Purushotham
- Tom Rhodes
- David Wylde

Background and Scope

Credibility methods have been widely applied to property, casualty, and health insurance for nearly a century. In contrast, credibility methods have seen limited use by life insurance companies and pension plans until recently. As a result, relatively few FSAs have experience in the application of credibility methods.

The recent adoption of principle-based reserves (PBR) by U.S. insurance regulators has created a need for life actuaries to become more familiar with credibility methods. Recent U.S. Treasury regulations have detailed a credibility method to be applied to pension plan mortality assumptions, raising the importance of credibility for pension actuaries.

This paper provides an overview of the most common credibility methods and a review of relevant credibility literature. Credibility methods are applied when setting rates used in practice areas such as pricing, experience rating, insurance valuation, and pension plan valuation.

Chapter 1: Overview of Credibility

In 1916, Albert H. Mowbray¹, a pioneer in the development of credibility, wrote, with some alteration to match the terminology used in this paper: “A credible rate is one for which the probability is high that it does not differ from the true, but unknown, underlying rate by more than a specified percentage.”

The credibility of an observed rate is enhanced by increased data: The more relevant data that can be brought to bear, the more credible the observed rate will be. To quantify the value of additional data, claim studies can calculate both the mean (i.e., the observed rate) and the standard deviation of claims. As the volume of data increases, the standard deviation of claims as a percentage of total claims will shrink. The smaller this percentage, the closer to full credibility. Full credibility is obtained when the observed rate has so much supporting data that any remaining error is inconsequential.

Don Behan, in a 2009 presentation², added clarity when he cautioned about using the term “credibility” with non-actuaries:

“We need to be careful in discussing the credibility of data to be clear that we are talking about statistics. Of course, the dictionary meaning of credible is “believable.” As actuaries, when we say that data has low credibility, we do not mean that it is not believable, but that we would assign less than 100% statistical weight in a weighted average. People who are unfamiliar with the statistical meaning of the word may incorrectly interpret our statement to mean that we don’t believe that the data is accurate. The term “statistical credibility” tends to make our meaning clearer.”

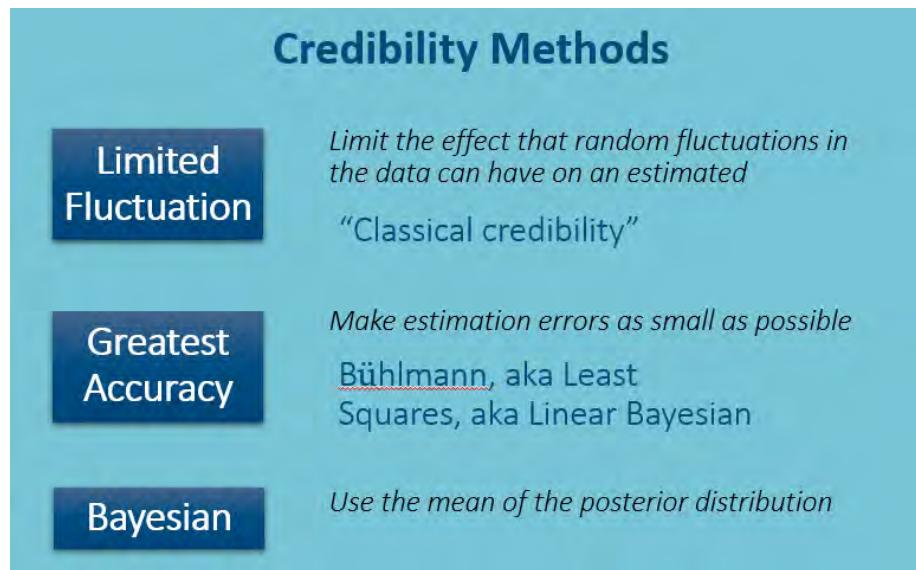
Credibility methods were originally developed to assist in the re-rating of premiums (i.e., calculating a new premium rate) in the property, casualty, and health insurance fields, starting a century ago. Credibility methods determined how much weight should be given to recent experience, with the remainder of the weight given to a prior premium rate or other data.

¹ [Proceedings of the Casualty Actuarial and Statistical Society of America, 1915-1916](#)

² June 18, 2009 presentation to the Southeastern Actuarial Conference

1.1 Credibility Methods

There are three major families of credibility methods used by actuaries today: Limited Fluctuation, Greatest Accuracy, and Bayesian. The following chart, inspired by a presentation by Paul Brehm³, shows how the various credibility methods are related:



The Limited Fluctuation method, which was the only credibility method used for insurance purposes until the 1960s, is sometimes dubbed “classical credibility.” The Greatest Accuracy method emerged in the 1960s and goes by at least the three different names shown in the box above. The Greatest Accuracy method is implemented through the use of models, such as Bühlmann-Straub, and estimation methods, such as Empirical Bayes.

The Bayesian method produces results based on combining a prior opinion regarding the unknown quantity (expressed as a probability distribution) and the observed data (expressed as a likelihood function, based on an assumed probability model).

³ Basic Concepts in Credibility, presented at the March 2006 Casualty Actuarial Society meeting in Minneapolis, MN

Each of the three credibility methods has its challenges:

1. The Limited Fluctuation method, though simple to implement and widely used, has theoretical shortcomings.
2. The Greatest Accuracy method, though more complex and less popular, is more theoretically-grounded.
3. The Bayesian method, which has not seen much use to date by actuaries, has considerable potential as software packages become more powerful and widely available.

Except where the Bayesian method is specifically mentioned, **all references to credibility and credibility methods from this point forward refer only to the Limited Fluctuation and Greatest Accuracy methods.**

1.2 General Formula for Credibility Methods

In practice, credibility methods are used for much more than just determining new premium rates. They also can be used to set mortality assumptions, disability assumptions, and various claim rates. In general, credibility methods derive an estimate of the true underlying rate as a linear compromise between an observed rate based on recent experience and either:

- the *Prior Rate*, for the Limited Fluctuation Method or
- the *Portfolio Rate*, for the Greatest Accuracy Method.

The formula for the *Credibility-Weighted Rate* is almost the same for the Limited Fluctuation and Greatest Accuracy methods:

- *Observed Rate* is a best estimate or a result from a recent experience study that applies to the risk group being priced or valued.
- *Prior Rate*, used by the Limited Fluctuation method, is the prior premium rate or a rate resulting from a much larger experience study. The Limited Fluctuation Method presumes the *Prior Rate* to be fully credible.
- *Portfolio Rate*, used by the Greatest Accuracy method, is the overall rate for a set of related risk groups, including the risk group to be estimated.
- Z is the credibility weight assigned to the *Observed Rate*, where Z is equal to 1 when full credibility is assigned to the *Observed Rate* and equal to 0 when no credibility is assigned to the *Observed Rate*.
- The remainder of the weight, $1 - Z$, sometimes called the complement of credibility, is the weight assigned to the *Prior Rate* or *Portfolio Rate*.
- The *Credibility-Weighted Rate* is calculated as the weighted average of the *Observed Rate* and either the *Prior Rate* (for Limited Fluctuation) or *Portfolio Rate* (for Greatest Accuracy):

For the Limited Fluctuation Method:

$$\text{Credibility-Weighted Rate} = (Z)\text{Observed Rate} + (1 - Z)\text{Prior Rate}.$$

For the Greatest Accuracy Method:

$$\text{Credibility-Weighted Rate} = (Z)\text{Observed Rate} + (1 - Z)\text{Portfolio Rate}.$$

Chapter 2: Limited Fluctuation Method

The Limited Fluctuation Method was so named because it allowed premiums to fluctuate from year to year based on experience, while limiting those fluctuations by giving less than full credibility to premiums based on limited data. In contrast, setting premium rates by giving full credibility to recent experience could be called the Full Fluctuation Method. While every credibility method serves to limit fluctuations, this method acquired its name because it was the first. The Limited Fluctuation Method, also known as Classical Credibility, is the most widely-used credibility method because it can be relatively simple to apply. Outside North America, this method is sometimes referred to as American Credibility.

The Limited Fluctuation Method has two components. It uses the formula introduced in the previous chapter to calculate the *Credibility-Weighted Rate*. This formula, repeated below, requires a credibility weight, Z , an *Observed Rate*, and a *Prior Rate*.

$$\text{Credibility-Weighted Rate} = (Z)\text{Observed Rate} + (1 - Z)\text{Prior Rate}.$$

The elements of this formula were defined in Chapter 1. The credibility weight, Z , is calculated using one of three formulas presented in this chapter: The square root formula, the asymptotic formula, and an approach based on confidence intervals.

2.1 Relevance of Prior Rate

The choice of the *Prior Rate* is very important for the Limited Fluctuation Method. A good choice for the *Prior Rate* often comes from answering the question “What estimate would be used if there were no experience data?”

A potentially large flaw for the Limited Fluctuation method is the applicability of the *Prior Rate*. For example, when an industry average is used as the *Prior Rate*, the inclusion of data from insurers with different target markets, different underwriting standards, and different distribution systems could result in a *Prior Rate* that detracts from, rather than enhances, the *Credibility-Weighted Rate*.

It may be possible to adjust the chosen *Prior Rate* to make it more applicable. Alternatively, it may be possible to reduce the scope of an industry or other large study to better match the business for which a *Credibility-Weighted Rate* is needed.

Another potential flaw is that the *Prior Rate* may be less than fully credible: Since the Limited Fluctuation method assumes that the *Prior Rate* has full credibility, it may receive more weight than it deserves.

2.2 Square Root Formula

Suppose we defined 90% credibility (and its complement, uncertainty of 10%) as a 90% probability that the *Observed Rate* is within plus or minus 5% of the true rate. If we increase observations by a factor of four, we can cut uncertainty in half, from 10% to 5%. This is because four times as many observations quadruples claims, while only doubling the standard deviation. More generally, if the number of observations is increased by X^2 , uncertainty is divided by X : in other words, uncertainty is inversely proportional to the square root of the number of observations. The square root formula is a result of this property of uncertainty and its complement, credibility. We will define the following two variables:

- N_{AC} , the actual observations (which could be measured as number of policies, number of claims, amount of claims, etc.) on which the *Observed Rate* is based, and
- N_{FC} , the observations required to achieve Full Credibility.

2.2.1 FULL CREDIBILITY

Full credibility can be defined as being achieved when the *Observed Rate* has so much supporting data that any remaining error is inconsequential.

Alternatively, full credibility could be defined more specifically as “less than a 2% chance that the percentage difference between the *Observed Rate* and the unknown true rate is more than 1%.” Assuming there is sufficient experience data so that the normal distribution can be applied, a 98% probability that the *Observed Rate* is within 1% of the true rate would require the following relationship:

$$\text{Standard Deviation} < \text{Observed Rate} * \frac{1\%}{2.32635},$$

where 2.32635 is the normal distribution Z-score associated with a 98% confidence interval, i.e., the number of standard deviations (plus and minus) required to include all but a 1% tail on either side of the normal distribution. This definition of full credibility would be rarely achieved.

Using a more practical sensibility, full credibility is more often defined as “less than a 10% chance that the difference between the *Observed Rate* and the unknown true rate is more than 5%.” Assuming there is sufficient experience data to apply the normal distribution, this would require the following relationship:

$$\text{Standard Deviation} < \text{Observed Rate} * \frac{5\%}{1.64485},$$

where 1.64485 is the normal distribution Z-score associated with a 90% confidence interval, i.e., including all but a 5% tail on each side of the normal distribution. This definition of full credibility is much more achievable, as demonstrated below, assuming the Poisson distribution yields $N_{FC} = 1,082$ claims.

All else being equal, a lower threshold for full credibility gives more weight to the *Observed Rate* and less weight to the *Prior Rate*. When the *Prior Rate* includes risks that differ materially from those being rated, a lower threshold for full credibility may be desirable.

Given N_{AC} and N_{FC} , the Square Root Formula determines the credibility weight, Z , as follows: If N_{AC} equals or exceeds N_{FC} , then full credibility has been achieved and Z is set equal to 1. Otherwise, Z is calculated using the following formula:

$$Z = \sqrt{\frac{N_{AC}}{N_{FC}}}.$$

When claims are of varying amounts, care must be taken to calculate N_{FC} using a standard deviation that reflects the varying claim amounts. Otherwise, credibility will be overstated, sometimes significantly so, as demonstrated in the numerical example in Section 2.5.

To demonstrate the use of the square root formula, suppose there were 270 observed claims and 1,082 claims were needed for full credibility. The square root formula would calculate the credibility weight as follows:

$$Z = \sqrt{\frac{270}{1,082}} = 0.4995.$$

2.2.2 DERIVATION OF THE NUMBER 1,082

The number 1,082 is the most commonly used measure of full credibility for the square root formula. For example, IRS Pension Regulation 10-4-2017 specifies that, for pension plans with at least 100 deaths and no more than 1,081 deaths, adjustments to the IRS mortality table are to be based on the Limited Fluctuation method's square root formula, with 1,082 deaths used as the threshold for full credibility. For pension plans with fewer than 100 deaths, the IRS mortality table must be used as is, with zero credibility (i.e., $Z = 0$) presumed.

The number 1,082 is based on full credibility being defined as a probability, P_{FC} , equal to 90% that the *Observed Rate* does not differ from the true rate by more than a ratio, R_{FC} , equal to 5%. The number 1,082 is derived as follows:

1. Assume that all claims are for the same amount.
2. Assume the Poisson distribution applies, which results in the mean and variance of claims equal to nq and the standard deviation equal to \sqrt{nq} , where n is the number of lives exposed and q is the observed claim rate. The Poisson distribution is not appropriate for mortality or lapse, as it allows for multiple claims per insured.
3. Use q as the best estimate of the underlying, but unknown, true claim rate.
4. Calculate probabilities using the normal distribution to approximate the Poisson distribution.

Using the definition of full credibility and the normal distribution z-value of 1.64485 for a 90% confidence interval, we set 1.64485 standard deviations equal to 5% of the expected number of claims (nq):

$$1.64485 \sigma = 1.64485 \sqrt{nq} = 5\% * nq.$$

Solving for nq , the number of claims required for full credibility, we have:

$$nq = \left(\frac{1.64485}{5\%} \right)^2 = 1,082.$$

2.2.3 COUNTERPART OF 1,082 FOR MORTALITY AND LAPSE RISKS

The standard deviation for mortality and lapse risks is calculated based on the binomial distribution, not the Poisson distribution, and is equal to \sqrt{npq} , where n is the number of lives exposed, q is the mortality or lapse rate, nq is the expected number of decrements, and p is equal to $1 - q$.

Using the binomial distribution, we can rewrite the relationship required for full credibility as:

$$1.64485 \sigma = 1.64485 \sqrt{npq} = 5\% * nq.$$

Solving for nq , the number of deaths or lapses required for full credibility, we have:

$$nq = p \left(\frac{1.64485}{5\%} \right)^2 = p 1,082.217.$$

Inspecting the above formula, we see that the use of 1,082 claims as a proxy for “full credibility defined as a 90% probability that the *Observed Rate* will be within 5% of the true rate” is accurate only for very small decrement rates where p is nearly equal to one. The larger the mortality or lapse rate, the fewer claims that are required for a given level of credibility. To illustrate, the following are needed to achieve full credibility as defined above:

- For $q = 0.01$, 1,071 decrements or 107,139 lives exposed.
- For $q = 0.5$, 541 decrements or a mere 1,082 lives exposed.

It might seem that a probability near 0.5 would be the hardest to estimate, since its variance is maximized. However, the credibility formula works with *relative* error. An acceptable error of 5% of $q = 0.01$ amounts to an acceptable error of just 0.0005. For $q = 0.5$, a 5% acceptable error amounts to 0.0250, which is a much easier threshold to achieve.

2.3 Asymptotic Formula

In his paper titled The Theory of Experience Rating, Albert W. Whitney⁴ developed the following asymptotic formula for calculating the credibility weight:

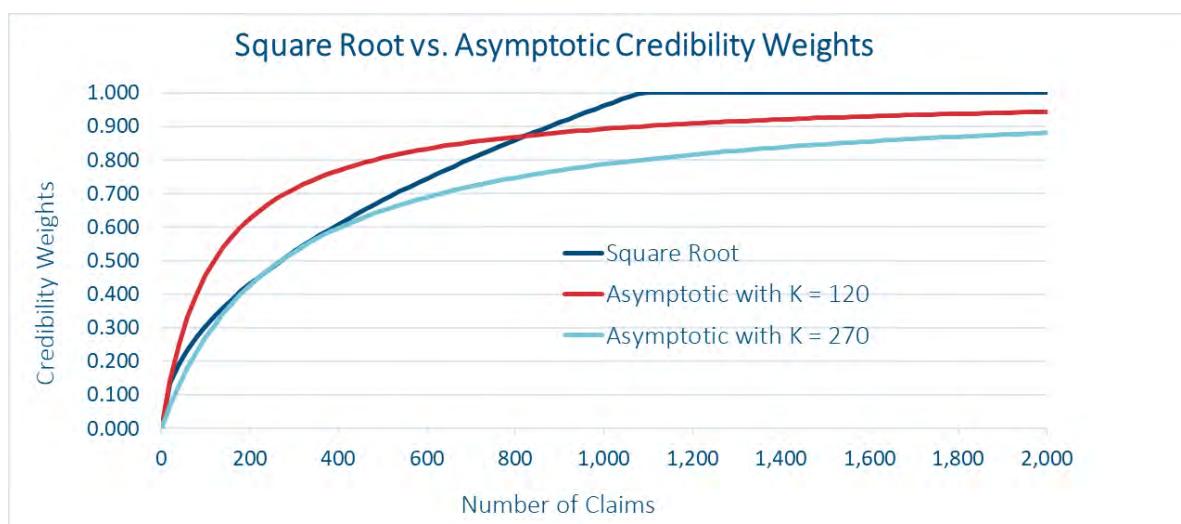
$$Z = \frac{N}{N+K},$$

where N is a measure of exposure, such as the number of actual claims, and K is a constant. This formula produces a credibility weight of $Z = 0$ for no experience data. The credibility weight then increases at a decelerating rate as experience data grows, asymptotically approaching $Z = 1$ as experience data increases to infinity. When used with the Limited Fluctuation method, K is determined by the actuary's judgment or set by precedence, such as by using a value of K that has been used for similar situations.

When K is set based on judgment, a simple approach is to solve for the value of K required to achieve a targeted level of partial credibility for a given number of claims. For example, K could be set so that 50% credibility is achieved when the number of claims is one-fourth of the 1,082 claims (i.e., 270 claims) often required for full credibility. This would result in $K = 270$. A rationalization for the choice of $K = 270$ is that the square root formula, when used with 1,082 claims as the measure of full credibility, also produces $Z = 0.5$ for 270 claims.

An alternative might be to set $K = 120$, which would produce $Z = 90\%$ when $N_{AC} = 1,082$ (which is the number of claims required to achieve a 90% probability of the *Observed Rate* being within plus or minus 5% of the underlying, but unknown, true rate).

The following graph compares results for the Square Root formula using $N_{FC} = 1,082$ to results for the Asymptotic Formula using both $K = 120$ and $K = 270$.



⁴ Proceedings of the Casualty Actuarial Society, 1918.

The square root formula and the asymptotic formula with $K = 270$ produce surprisingly similar results from 100 to 400 claims. This is because both produce $Z = 0.5$ at 270 claims. Beyond 400 claims, the square root formula increases almost linearly to achieve full credibility at 1,082 claims. In contrast, the asymptotic formula with $K = 120$ produces a value of $Z = 0.9$ for 1,082 claims. This may be more reasonable since the only perfect estimate would be one based on an infinite sample. Any finite sample should theoretically receive less than full credibility.

2.4 Credibility Weights Based on Confidence Intervals

Credibility could be calculated as the probability that the observed mean is within a tight confidence interval, such as plus or minus 1%, of the true mean. Such a probability can be estimated based on the observed mean, the observed standard deviation, and an assumed distribution. For large numbers of claims, the normal distribution often produces satisfactory results. For smaller numbers of claims, other distributions may be more appropriate, such as the binomial distribution for deaths and lapses, and the Poisson distribution for most other decrements.

We will use the following notation:

- μ^T will represent the true but unknown mean,
- σ^T will represent the true but unknown standard deviation,
- μ will represent the observed mean,
- σ will represent the observed standard deviation, and
- r will represent the ratio that specifies the tight confidence interval, $\mu^T(1 \pm r)$, around the true mean.

The credibility weight, Z , can then be calculated as the probability that a confidence interval, running from $\mu^T(1 - r)$ to $\mu^T(1 + r)$, encompasses μ :

$$Z = \Pr[\mu^T(1 - r) \leq \mu \leq \mu^T(1 + r)].$$

To illustrate how this would work, we will assume the normal distribution applies. Adjusting the above formula to use the standard normal distribution's mean of zero and standard deviation of one (accomplished by subtracting μ^T from each term and then dividing by σ^T), we have:

$$Z = \Pr\left[-r \frac{\mu^T}{\sigma^T} \leq \frac{(\mu - \mu^T)}{\sigma^T} \leq r \frac{\mu^T}{\sigma^T}\right].$$

Assuming that $\frac{(\mu - \mu^T)}{\sigma^T}$ is normally distributed with a mean of zero and a standard deviation of one, Z can be calculated as the cumulative normal distribution from $-r \frac{\mu^T}{\sigma^T}$ to $r \frac{\mu^T}{\sigma^T}$. This result is a function of two variables: the confidence interval ratio, $\pm r$, and the ratio of the true mean to the true standard deviation, $\frac{\mu^T}{\sigma^T}$. Since both components of this ratio are unknown, we will assume that the ratio can be fairly approximated by substituting $\frac{\mu}{\sigma}$ for $\frac{\mu^T}{\sigma^T}$.

When the mean is much larger than the standard deviation, the resulting high ratio increases credibility. Conversely, when the standard deviation is much larger than the mean, the resulting low ratio lowers credibility.

The following table illustrates the credibility weight, Z , resulting from various combinations of confidence intervals (1%, 2.5%, and 5%) and ratios of mean to standard deviation (100, 50, 20, 10, 5, 2, 1, 0.5, 0.2, 0.1).

CREDIBILITY WEIGHTS				
Ratio of Mean to Std Dev.	Confidence Interval			
	$\pm 1.0\%$	$\pm 2.5\%$	$\pm 5.0\%$	
100.0	68.3%	98.8%	100.0%	
50.0	38.3%	78.9%	98.8%	
20.0	15.9%	38.3%	68.3%	
10.0	8.0%	19.7%	38.3%	
5.0	4.0%	9.9%	19.7%	
2.0	1.6%	4.0%	8.0%	
1.0	0.8%	2.0%	4.0%	
0.5	0.4%	1.0%	2.0%	
0.2	0.2%	0.4%	0.8%	
0.1	0.1%	0.2%	0.4%	

To illustrate how the above table was populated, here are the steps used to determine the credibility weight of 68.3% for a mean to standard deviation ratio of 100 and a confidence interval of $\pm 1\%$:

1. The mean to standard deviation ratio was multiplied by the confidence interval to obtain normal distribution z-values of 1.0 and -1.0, the endpoints of the desired cumulative distribution.
2. The cumulative normal distribution associated with a z-value of 1.0 was looked up in a table and found to be 0.84134.
3. 0.5 was subtracted from 0.84134 to determine the probability of the true mean lying between z-values of 0 and 1.0, resulting in a probability of 34.134%.

4. Since the normal distribution is symmetric, the probability of 34.134% was doubled to obtain a probability of 68.3% for the true mean lying between z-values of -1.0 and 1.0.

2.5 Mortality Study Example

We will focus on a tiny part of a life insurance company's mortality study: The credibility of the ultimate mortality rate for male age group 63 to 67 and policy years 25+. The company has chosen to use a *Prior Rate* equal to the age 65 rate from the 2001 VBT Ultimate Table for Male, Composite risks, which is equal to 0.01588.

2.5.1 MORTALITY STUDY DATA AND RESULTS

For simplicity, the example illustrates only four possible policy sizes. The following table shows mortality study results, consisting of exposure and deaths by both count and amount, for each policy size and in total:

MORTALITY STUDY DATA				
Policy Size	Count Exposed	Count of Deaths	Amount Exposed	Amount of Deaths
\$ 50,000	12,800	210	\$ 640,000,000	\$ 10,500,000
\$ 100,000	3,200	49	\$ 320,000,000	\$ 4,900,000
\$ 250,000	800	11	\$ 200,000,000	\$ 2,750,000
\$ 500,000	200	3	\$ 100,000,000	\$ 1,500,000
Total	17,000	273	\$ 1,260,000,000	\$ 19,650,000

The next table shows the following key results from the mortality study:

- Mortality rates by count, q_{cnt} , and amount, q_{amt} , are calculated as total deaths divided by total exposed.
- Variance by count is calculated as $(count\ exposed)(q_{cnt})(1 - q_{cnt})$.
- Variance by amount is calculated separately for each policy size as $(count\ exposed)(policy\ size)^2(q_{amt})(1 - q_{amt})$.

Policy Size	MORTALITY STUDY RESULTS			
	Mortality Rates		Variance	
	by Count	by Amount	by Count	by Amount
\$ 50,000	0.01641	0.01641	202.252	491,265
\$ 100,000	0.01531	0.01531	50.563	491,265
\$ 250,000	0.01375	0.01375	12.641	767,601
\$ 500,000	0.01500	0.01500	3.160	767,601
Total	0.01606	0.01560	268.616	2,517,732

While it is not necessary to calculate total variance by count separately for each policy size, doing so allows for an interesting comparison to total variance by amount: Variance by count decreases dramatically as policy size increases, due to declining exposure, while total variance by amount is flat to increasing by policy size. Both of these trends are not unusual.

2.5.2 DETERMINATION OF THRESHOLD FOR FULL CREDIBILITY

Three types of standard deviations are shown in the next table:

- Standard deviation calculated as the square root of total variance.
- Standard deviation as a percentage of deaths, calculated as standard deviation divided by total death count or total death amount.
- The **full credibility standard deviation** is the standard deviation required to achieve full credibility. It is based on $P = 90\%$ and $r = 5\%$ and expressed as a percentage of deaths. It is calculated as 5% divided by 1.64485 , which is the normal distribution z-value for a 90% confidence interval.

A rate whose standard deviation, as a percentage of deaths, does not exceed the full credibility standard deviation is given full credibility. The larger the standard deviation, the lower the credibility, all else being equal.

STANDARD DEVIATIONS		by Count	by Amount
Standard Deviation (= $\sqrt{\text{Variance}}$)		16.390	\$ 1,586,736
Std Dev as % of Deaths		6.003%	8.075%
Full Cred Std Dev as % of Deaths		3.040%	3.040%

In the preceding table, the standard deviation as a percentage of death count is almost twice that needed for full credibility, which means that almost four times as many deaths would be required to achieve full credibility. The standard deviation as a percentage of death amount is 2.66 times that needed for full credibility, which means that slightly more than seven times as many deaths would be required to achieve full credibility. Note there is a significant difference between the standard deviations by count and amount as a percentage of deaths. This illustrates the significant overstatement of credibility that can occur when count-based credibility measures are applied to amount-based rates.

In the following table, we calculate the **Study-to-Full Cred Ratio** as A divided by B , where:

A is the study's standard deviation as a percentage of deaths, and

B is the full credibility standard deviation as a percentage of deaths.

It requires four times as many deaths to halve the standard deviation as a percentage of deaths. More generally, to decrease the study's standard deviation to match the target standard deviation would require deaths to increase by the Study to Full Cred Ratio squared. In the last line of the table, study deaths are multiplied by the Study to Full Cred Ratio squared to determine the count and amount of deaths required to achieve full credibility.

DEATHS REQUIRED FOR FULL CREDIBILITY		by Count	by Amount
Study to Full Cred Ratio		1.975	2.656
Study to Full Cred Ratio squared		3.901	7.057
Study Death Count and Amount		273	19,650,000
Full Cred Death Count and Amount		1,065	138,663,294

Because of the relatively larger standard deviation by amount, seven times more dollars of death claims would be required for full credibility, compared to only four times as many deaths.

2.5.3 CALCULATION OF CREDIBILITY-WEIGHTED RATE

Having calculated the death count and amount required for full credibility, we can apply the Limited Fluctuation Method's square root formula to calculate Z as the square root of the ratio of actual deaths to full credibility deaths. The *Observed Rate* is then weighted by Z , while the *Prior Rate* is weighted by $1 - Z$ to obtain the *Credibility-Weighted Rate*.

The following table shows an example of the square root formula applied to a rate based on counts and to another rate based on amounts. The value of Z in the "by Count" column is close to 0.5, reflecting a fair degree of credibility. In contrast, the value of Z in the "by Amount" column is very nearly equal to zero, reflecting extremely low credibility due to the much higher standard deviation associated with the wide distribution of policy sizes.

SQUARE ROOT FORMULA		by Count	by Amount
Observed Rate		0.01606	0.01560
Prior Rate		0.01588	0.01588
Actual Deaths / Full Cred Deaths		0.25638	0.14171
$Z = \sqrt{\text{Actual to Full Cred}}$		0.50634	0.37644
Credibility-Weighted Rate		0.01597	0.01577

For the asymptotic formula, we know N , the observed death count or amount, but we need to establish a value for K . We will do so by arbitrarily setting it equal to 50% of the death count and death amount needed for full credibility, as shown in the next table. Once Z is calculated, the *Credibility-Weighted Rate* can be calculated using the general credibility formula.

ASYMPTOTIC FORMULA		by Count	by Amount
Observed Rate		0.01606	0.01560
Prior Rate		0.01588	0.01588
$N = \text{Observed count or amount}$		273	19,650,000
$K = 50\% \text{ of Full Cred Deaths}$		532.4	69,331,647
$Z = N / (N + K)$		0.33895	0.22083
Credibility-Weighted Rate		0.01594	0.01582

2.6 Credibility Applied to Mortality Ratios

The preceding example was based on mortality rates, which makes for a simpler example. However, the results of many mortality studies are expressed as mortality ratios of the experience under study to a standard industry table. For the application of the Limited Fluctuation method to the credibility of mortality ratios, please refer to one of the following:

- Pages 19-21 of the Credibility Practice Note dated July 2008, issued by the American Academy of Actuaries.
- Pages I.4 to I.10, of Credibility Theory Practices dated December 2009, issued by the Society of Actuaries. This includes 4 pages of numerical results.

2.7 Strengths and Weaknesses

STRENGTHS OF THE LIMITED FLUCTUATION METHOD

- The Limited Fluctuation Method is relatively simple and easy to apply: In situations where all claims are for the same amount, only the number of claims is needed to determine full credibility or calculate the credibility weight for partial credibility.
- The method is useful for experience rating when there is an appropriate *Prior Rate*.
- The method's results might be improved by calculating the observed claim's standard deviation as a percentage of claims and using it to calculate the credibility weight as a probability, as described in Section 2.4.
- The only data needed is from the group of risks being assessed, in addition to a *Prior Rate*.

WEAKNESSES OF THE LIMITED FLUCTUATION METHOD

- The derivations of the formulas are arbitrary.
- The inputs to the formulas are arbitrary.
- The same weight, $1 - Z$, is given to the *Prior Rate*, regardless of the analyst's view of the accuracy of the *Prior Rate*.
- When the rate is based on claim amounts rather than claim counts, the use of claim counts as inputs can significantly overstate credibility.
- Results may be distorted when there is no appropriate choice for the *Prior Rate*, such as when there is no credible experience for a new type of insurance coverage.
- The AAA's Credibility Practice Note issued in July 2008 cited an additional weakness: Most derivations for the method use variance as a measure of quality. However, credibility estimates are biased, making mean square error (which is the measure used with the Greatest Accuracy method) a more reasonable choice.
- The Square Root formula reaches full credibility prematurely, rather than asymptotically approaching $Z = 1$ without ever reaching it.

Chapter 3: Greatest Accuracy Method

3.1 Bayesian Credibility, Stepping Stone to Greatest Accuracy

To paraphrase Enrique de Alba⁵, “Very frequently in actuarial science, one has considerable prior information in the form of global or industry-wide experience or in the form of tables. In this respect, it is surprising that Bayesian methods have not been used more intensively, as there is a wealth of ‘objective’ prior information available to the actuary.” While Bayesian credibility makes use of such prior information, the assumed distribution of results is based on the analyst’s judgment.

According to Udi Makov⁶, “Bayesian ideas were introduced into actuarial science in the late 1960s in the form of empirical credibility methods for premium setting. The advance of the Bayesian methodology was slow due to its subjective nature and to the computational difficulties associated with the full Bayesian analysis.” Due to advancements in computer hardware and software, Bayesian credibility is no longer difficult to apply. Previously, certain combinations of distributions, such as normal distributions, were used to simplify the mathematics.

The American Academy of Actuaries’ July 2008 Credibility Practice Note reported that “With the increased power of 21st-century computing equipment, advances in statistical algorithms (e.g., the EM algorithm and Markov chain Monte Carlo methods) to implement the Bayesian approach, and widely-available software that performs Bayesian inference (i.e., WinBUGS⁵), a wider class of problems is becoming susceptible to solution via the Bayesian approach.”

From a Bayesian advocate’s point of view, Bayesian credibility produces the best estimate of the true underlying rate, when evaluated using expected squared error. However, unlike the Greatest Accuracy method, the Bayesian method produces a result that will not necessarily lie between the observed mean for the risk group and the overall observed mean for a portfolio of related risk groups.

3.2 Overview of Greatest Accuracy

Bayesian credibility served as the starting point for Bühlmann’s development of the Greatest Accuracy method. The Greatest Accuracy credibility method goes by several aliases: Bühlmann credibility, Least Squares credibility, and Linear Bayesian credibility. These aliases resulted from Bühlmann’s development of a credibility method that was the **least squares linear approximation** to **Bayesian** credibility.

The American Academy of Actuaries’ July 2008 Credibility Practice Note states that, compared to Bayesian credibility, Greatest Accuracy credibility typically “requires less mathematical skill,

⁵ [Bayesian Claims Reserving](#), Encyclopedia of Actuarial Science

⁶ [Principal Applications of Bayesian Methods in Actuarial Science](#), North American Actuarial Journal, October 2001

requires fewer computational resources, and does not require the selection of a prior distribution, which can sometimes require difficult judgments.”

According to Vincent Goulet⁷, the Limited Fluctuation Method’s purpose is “to incorporate into the premium as much individual experience as possible, while still keeping the premium sufficiently stable.” Goulet goes on to state that “When credibility is used to find the most precise estimate of an insured’s pure risk premium, one must turn to Greatest Accuracy credibility methods.”

While Bühlmann expressed strong support for the Bayesian credibility method, he felt that, whenever possible, assumptions regarding the underlying distributions should be based on experience data rather than subjective judgment. This led him to develop and publish a new method in 1967 that produced credibility-weighted rates that were the best linear approximation to the Bayesian method’s estimate. The credibility weight, Z , was calculated to minimize the expected value of the square of the differences between the Bühlmann-weighted estimate and the Bayesian predictive mean.

3.3 Greatest Accuracy Formulas

The Greatest Accuracy method assumes that each risk group in the portfolio produces results according to an unknown probability distribution. A second assumption is that the probability distributions for the various risk groups are distributed according to another unknown probability distribution. The method then combines these assumptions with data from the portfolio of risk groups to develop the best linear estimate of the true rates.

The Greatest Accuracy method’s *Credibility-Weighted Rate* formula is similar to that used by the Limited Fluctuation method, with *Portfolio Rate*, the *Observed Rate* for all risk groups combined, used in place of *Prior Rate*:

$$\text{Credibility-Weighted Rate} = (Z)\text{Observed Rate} + (1 - Z)\text{Portfolio Rate}.$$

The Greatest Accuracy method uses means and variances from a portfolio of risk groups to calculate a risk group’s *Credibility-Weighted Rate*. The portfolio would consist of related risk groups, such as groups of similar risks from different geographical areas, different occupation groups, or different insurance companies.

The credibility weight for each risk group is influenced by a combination of the risk group’s results and the overall portfolio’s results:

- When a risk group has relatively large exposure or number of claims, more weight will be given to its observed mean. Conversely, when a risk group has a relatively small exposure or number of claims, less weight will be given to its observed mean.

⁷ Principles and Application of Credibility Theory, Vincent Goulet, 1998

- When means vary significantly between risk groups, more weight will be given to each risk group's observed mean. Conversely, when means vary only slightly between risk groups, more weight will be given to the overall portfolio's mean.

The Greatest Accuracy method results in the following formula, sometimes called the Bühlmann Credibility Formula, for the credibility weight, Z :

$$Z = \frac{N}{N+K},$$

where K is a constant and N is exposure or some other measure of sample size, such as number of claims or earned premium.

The constant K , sometimes called the Bühlmann credibility constant, is solved-for mathematically, with the following result:

$$K = \frac{v}{a} = \frac{\text{Expected Value of the Process Variance}}{\text{Variance of the Hypothetical Means}} = \frac{E[\text{Var}[X]]}{\text{Var}[E[X]]}.$$

The numerator, $v = \text{Expected Value of the Process Variance} = E[\text{Var}[X]]$, is a measure of the average variance within the portfolio's risk groups.

The denominator, $a = \text{Variance of the Hypothetical Means} = \text{Var}[E[X]]$, is a measure of the variance between risk group means.

For those many years removed from the study of probability, $E[\text{Var}[X]]$ and $\text{Var}[E[X]]$, may be somewhat inscrutable. If so, please refer to Appendix A for a numerical example that illustrates how to calculate $E[\text{Var}[X]]$ and $\text{Var}[E[X]]$ when underlying distributions are precisely known.

When experience data are used to calculate $E[\text{Var}[X]]$ and $\text{Var}[E[X]]$, the formulas are not quite as straightforward as those based on known underlying distributions. This is because of the effect of random fluctuations on the variance of hypothetical means, which add a bias that must be removed. The correction of that bias is the gist of the Bühlmann-Straub Empirical Bayes estimation formula.

Restating the previous formula for Z by substituting $\frac{v}{a}$ for K , we have:

$$Z = \frac{N}{N+\frac{v}{a}}.$$

The following characteristics can be observed for the above formula:

- As N , exposure or some other measure of sample size, increases, so does credibility, as measured by Z , approaching unity as N approaches infinity.
- A high value of ν , a measure of the average variance for the portfolio's risk groups, indicates that observations within risk groups are highly variable, which implies less credible risk group means, resulting in a lower value of Z .
- A large value of a , a measure of the overall variance between risk group means, indicates that risk group means are highly variable. This results in a larger value of Z , with more weight given to each risk group's mean and less weight given to the portfolio mean.
- Conversely, a small value of a indicates there is little difference between risk group means. This results in a smaller value of Z , so that less weight is given to each risk group's mean and more weight is given to the portfolio mean.

To clarify the last two points, consider an extreme case with $a = 0$, which would indicate that all risk groups have the same mean. In that case, the best estimate of the mean for a particular group is the portfolio mean (hence $Z= 0$), since there is no point in estimating each group separately.

3.4 Bühlmann-Straub Model

In 1970, Bühlmann and Straub extended Bühlmann credibility (also known as Least Squares and Linear Bayesian credibility). The Bühlmann-Straub model allows for variations in exposure and claim size.

The mathematics underlying this method are beyond the scope of this paper. However, Hardy and Panjer authored a paper titled [A Credibility Approach to Mortality Risk](#) that describes their application of Bühlmann-Straub credibility to amount-based mortality ratios by company. Also, Bühlmann and Gilser created [A Course in Credibility Theory and its Applications](#) to teach how to apply Bühlmann-Straub credibility, cited in the next chapter.

With Bühlmann credibility, the observed random variable is a single observation, such as annual claims for an insured. The Bühlmann-Straub model supports an observed random variable with both a claim number and its associated exposure. This is quite useful for situations where exposures routinely vary, such as in mortality and lapse studies.

3.5 Estimating the Components of K

When using Greatest Accuracy credibility, Empirical Bayesian estimates are commonly used to determine the components of K , equal to $\frac{v}{a}$, where:

$$v = \text{Expected Value of the Process Variance} \text{ and}$$

$$a = \text{Variance of the Hypothetical Means.}$$

Another approach to estimating these two parameters is to minimize the mean-squared error by setting two partial derivatives equal to zero, thereby created two equations with two unknowns. This approach is illustrated in Appendix B: Minimizing Mean-Squared Error.

3.6 Strengths and Weaknesses

The following strengths and weaknesses of the Greatest Accuracy credibility method were largely obtained from the American Academy of Actuaries' Credibility Practice Note, July 2008, and edited to fit the terms used in this paper.

Strengths of the Greatest Accuracy Method:

- All aspects of the Greatest Accuracy method are clearly spelled out. Any approximations or assumptions made are clearly stated.
- Once assumptions and objectives are in place, an exact answer can be obtained by applying basic probability principles.
- There is a clear objective function: minimizing the mean-squared error.
- There are no arbitrary choices that are unrelated to the observed random variables.

Weaknesses of the Greatest Accuracy Method:

- The Greatest Accuracy method requires a portfolio of data from comparable risk groups to estimate the method's parameters and *Portfolio Rate*. For example, a company cannot use the Greatest Accuracy method to determine the credibility of its own mortality experience unless it has access to similar data from other companies.
- If the portfolio of data includes risk groups that are not sufficiently similar to the risk group whose *Credibility-Weighted Rate* is to be calculated, the *Portfolio Rate* may serve to distort rather than improve the estimate. This is analogous to problems with the suitability of the *Prior Rate* used by the Limited Fluctuation credibility method.
- The Greatest Accuracy method may produce a poor approximation when the random variable has a heavy tail.
- The use of Empirical Bayesian estimates can result in a negative value for the variance of observed means, which is nonsensical.

Chapter 4: Recommended Credibility Readings

The first three sections of this chapter describe three highly-recommended sources of basic credibility knowledge for life, health, and pension actuaries. The fourth section lists additional sources relevant to life, health, and pension business. The last section lists additional sources from the casualty actuarial field, to which we owe the development of today's credibility methods.

4.1 Foundations of Casualty Actuarial Science

For a primer on the concepts underlying credibility methods, as well as the most-used credibility methods, please refer to Chapter 8, Credibility, by Howard C. Mahler and Curtis Gary Dean. This is part of the Casualty Actuarial Society's textbook titled Foundations of Casualty Actuarial Science. While written for casualty actuaries, the credibility chapter is peppered with good explanations and examples that are accessible for life actuaries. A clever example involving a mixture of 4-sided, 6-sided, and 8-sided dice provides practical insight into the mathematics underlying Bühlmann credibility. The chapter also contains several tables that tie credibility to probabilities arising from the normal and Poisson distributions. Finally, it has exercises (and solutions) at the end of each section to test your understanding before proceeding to the next section.

4.2 Credibility Practice Note, American Academy of Actuaries, July 2008

This practice note is primarily a collection of examples that illustrate the application of credibility to a number of different situations. The practice note includes the following:

- A summary of U.S. regulations and regulatory practices that reference credibility.
- A brief but thoughtful introduction to credibility.
- Limited Fluctuation credibility applied to the mortality rate for a group of 1,000 lives.
- An introduction to Greatest Accuracy credibility with a good discussion of challenges and limitations. Formulas are introduced with no derivations. A mortality rate example involves two groups of lives. Calculations are shown with minimal explanation.
- Limited Fluctuation credibility applied to mortality ratios.
- A discussion of credibility considerations for a reinsurer setting mortality assumptions for a client company.
- Development of a credibility-weighted mortality ratio using a method documented in a 2002 Education Note published by the Canadian Institute of Actuaries.
- Limited Fluctuation credibility used to calculate adjusted lapse rates that are a blend of observed lapse rates and a table of expected lapse rates. A good summary of strengths and weaknesses is included.
- Credibility applied to a very simple example of one million policies with no claims. When compared to the two possible Limited Fluctuation outcomes, the Bayesian approach makes much more sense.

- Credibility applied to 5,000 policies that experience 1.6 claims per thousand policies vs. a regulatory *prima facie* rate of 4.0 claims per thousand. The regulator uses a Bayesian approach to argue for a 33% decrease in premium rate, while the insurer uses Limited Fluctuation to justify a 5% decrease in premium rate. A good commentary explores the differences in the two approaches.
- A brief four-page history of the evolution of credibility theory.
- A two-page bibliography of recommended readings.
- Credibility-related regulatory guidance from six states.

4.3 Credibility Theory Practices, SOA Report, December 2009

This SOA report applies the Limited Fluctuation and Greatest Accuracy credibility methods to mortality and lapse ratios for a portfolio of 10 life insurers. Results are produced for rates based on both counts and amounts. Mortality results for each company are calculated as A/E ratios to the Nonsmoker 2001 VBT mortality table. Lapse results for each company are calculated as a percentage of the lapse rates for all companies combined. The development of formulas is well documented; however, intermediate results are not included.

CREDIBILITY SURVEY REPORT

The report included a survey of the use of credibility by U.S. life insurance companies. Out of 190 company responses to the survey, only 19 provided enough data to be included in the survey. The major conclusion was that credibility methods were not widely used by U.S. life insurers.

For those who used credibility methods, they were more often applied to mortality (90%) than to lapse and expenses (10%). Large life and annuity companies were more likely than others to use credibility methods. Most respondents could not name a credibility method that they used. Many quoted “actuarial judgment” as the method used.

Survey conclusions:

1. Life insurers had a low level of awareness of or interest in credibility methods.
2. There was a lack of credibility training and education.
3. There may be a need to conduct more research, develop practice notes, and write articles to further the level of interest in and knowledge of credibility methods.
4. Any future survey should include consultants and actuaries working in non-insurance fields.
5. Two factors contributed to the low number of responses to the survey:
 - The large number of questions asked and
 - Inability to identify the proper person at each company to receive the survey.

4.4 Other Recommended Readings for Life, Health, and Pensions

The following sources, listed in alphabetical order, provide additional insights for the application of credibility methods to life, health, and pension business:

- Application of Credibility Theory to Group Life Pricing, by Manuel Tschupp
 - This paper extends the Bühlmann-Straub credibility method to include the following novel additions:
 - A multidimensional approach was used to combine “by count” and “by amount” data into a single model.
 - Another multidimensional approach was used to mute the following unwanted fluctuations: Experience study results based on a running average of, say, five years of data, experience unwanted fluctuations when a year with unusual results is no longer included in the running average.
 - The paper is highly technical; it may require a higher proficiency with probability and statistics than most actuaries have.
- Credibility Analysis for Mortality Experience Studies, a very readable three-part series of articles by David Wylde published in The Messenger (Transamerica Reinsurance’s Risk Management Newsletter). Topics included the following:
 - Mortality as a Binomial Process
 - σ/μ (i.e., standard deviation divided by the mean) as a measure of credibility
 - Connection between the confidence level and claim count
 - Sources of fully credible data
- A Credibility Approach to Mortality Risk, by M.R. Hardy and H.H. Panjer
 - This paper applies Bühlmann-Straub credibility to develop credibility-weighted mortality ratios for individual life business of Canadian life insurance companies.
 - The paper ranges from quite readable to very technical.
- Credibility in PBR – Session 73 at the 2016 SOA Annual Meeting, by Steven Ekblad
 - A high-level, easy-to-follow review of the application of credibility for PBR
 - Where and how credibility comes into play – slides 3-5
 - Limited Fluctuation credibility in VM-20 – slides 6-7
 - Greatest Accuracy credibility with Empirical Bayesian estimates in VM-20
 - Comparison of Limited Fluctuation and Greatest Accuracy credibility
- Expected Mortality: Fully Underwritten Canadian Individual Life Insurance Policies, Committee on Life Insurance Financial Reporting, Canadian Institute of Actuaries, July 2002
 - This paper provides an instructive overview on how to apply credibility criteria to blend industry and company-specific experience data to construct expected mortality assumptions.

- The recommended approach to credibility is the Limited Fluctuation method with full credibility defined as a 90% probability that the observed rate is within 3% of the underlying true rate. This results in 3,007 death claims being required for full credibility.
- A **Normalized Method** is presented that uses the credibility and actual-to-expected ratio of subsets of a company's business to create "blended ratios" for these subsets. The blended ratios are then adjusted, in order, to reproduce the overall company ratio.
- Experience Studies – Session 41 at the 2012 SOA Annual Meeting, by Steven Ekblad
 - Easy-to-follow and thoughtful presentation that includes the following:
 - Results by count vs. amount of claims – slides 11-13
 - Monte Carlo vs. Panjer Methods – slides 14-15
 - Confidence Intervals – slides 16-17
 - Credibility with "f factor" to adjust from count to amount of claims – slides 18-25
- Group LTD Credibility Study, Results from Stage 1, by Paul Correia and Tasha Khan, sponsored by the SOA Group LTD Experience Committee
 - Someone familiar with Group LTD may find this study quite readable and informative.
 - The primary focus was on LTD claim credibility for pricing and underwriting purposes. Because of group dynamics, which result in non-independent claims, traditional credibility methods may not be applicable to group LTD.
 - Instead of assuming a true, unchanging underlying rate, credibility for group LTD requires additional considerations for claim rates that drift or change over time due to changes in unemployment levels, company practices, and modeling bias. The study, therefore, compared future experience to historical experience by computing both correlation coefficients and relative errors.
 - The paper computed correlation coefficients for employee group sizes ranging from 0-99 to 50,000+ employees. Many different combinations of variables were explored. For example, two-year experience periods were examined for look-back periods ranging from one year to five years. A graph shows that results for the five look-back periods were highly correlated.
 - Likewise, many different combinations of variables were examined for relative errors.

- Long-Term Care Credibility Monograph, by the eponymous Work Group of the American Academy of Actuaries
 - Focuses on the challenges of misestimation and volatility of LTC results
 - Includes a brief history and overview of Limited Fluctuation and Greatest Accuracy credibility methods
 - Portions of the monograph could prove useful for business types other than LTC
- NAIC Valuation Manual, January 1, 2018 Edition
 - VM-20: PBR for Life Products
 - Section 9.A.6a on page 44 describes to which risk factors credibility methods must be applied.
 - On pages 50-56, Section 9.C.4 spells out credibility-related procedures for mortality assumptions in great length. Not principle-based.
 - Pages 59-60 contain tables of grading periods to be used based on the credibility of company data.
 - VM-21: PBR for Variable Annuities
 - Credibility is mentioned repeatedly on pages 72-75
 - On pages 76-77, detailed guidance is spelled out for credibility adjustments.
 - VM-31: PBR Actuarial Report Requirements
 - Guidance for documenting credibility on page 5
 - Mentions of credibility on pages 6 and 12
 - VM-50: Experience Reporting Requirements
 - Guidance related to statistical credibility theory on page 44
 - Mentions of credibility on pages 46-47
- Selecting Mortality Tables: A Credibility Approach, by Gavin Benjamin, October 2008
 - Focuses on development of mortality assumptions for pension plans
 - Examines mortality rates based on lives (i.e., counts) vs. amounts.
 - Requirements for full credibility borrow from the Limited Fluctuation method
 - Applies credibility to observed rates vs. an existing mortality table
- Useful Tools for Assessing Claims Fluctuation, The Messenger, article by David Wylde
 - Demonstrates that a compound Poisson distribution can be used to model the distribution of claims that vary by amount

4.5 Other Recommended Readings from Casualty Actuarial Practice

The following sources, listed in alphabetical order, provide additional insights into the application of credibility methods used by casualty actuaries:

Excerpts from the following casualty actuarial sources provide straightforward explanations and analyses for one or more credibility methods:

- Basic Concepts in Credibility, presentation by Paul Brehm from a 2006 CAS Seminar on Ratemaking
 - Derivation of the Limited Fluctuation method – slides 10-14
 - Weaknesses of the Limited Fluctuation method – slide 22
 - Pictorial Introduction to the Greatest Accuracy method – slides 24-26
 - Derivation of Greatest Accuracy formulas – slides 27-29
 - Strengths and Weaknesses of the Greatest Accuracy method – slide 30
- The Complement of Credibility, by Joseph A. Boor
 - Considerations when developing or choosing the “complement of credibility,” referred to as the “Prior Rate” in this paper – pages 325-328
- A Course in Credibility Theory and its Applications, by H. Bühlmann and A. Gilser, published by Springer, 2005.
 - This is cited as an introduction to Bühlmann-Straub credibility.
 - Not reviewed because it requires a purchase; its level of understandability is unknown but it is apt to be quite technical.
- Credibility, An American Idea, by Charles C. Hewitt, Jr.
 - A brief history of credibility practices – pages 3-6
- Credibility, Practical Applications, by Howard Mahler
 - Helpful rules for applying credibility – pages 1-2
- Credibility Theory for Dummies
 - Five recommended pages: Definitely not for dummies – pages 623-627
- Credibility Theory and GLM, by Esbjorn Ohlsson and Bjorn Johansson
 - This paper presents an approach to multi-dimensional modeling using a combination of random and fixed effects.
 - The paper explores the relation between credibility and GLMs and derives predictors of the random effects by minimizing mean-squared errors.
 - Not a simple read.
- An Introduction to Basic Credibility, by Howard C. Mahler
 - Transcript from a March 1996 CAS Ratemaking Seminar
 - Although the focus is on rates for workers compensation, the examples help with understanding credibility basics.
 - Exhibits 1-22 can be found at the end of the presentation.
 - Exhibits 2, 3, 5, 7, 9, and 10 are pictorial distributions of loss experience that likely used different colors for different risk groups. Without colors, they are hard to interpret.
 - Exhibit 13 presents credibility-related rules that may be quite helpful.

- An Introduction to Credibility, by Curtis Gary Dean
 - P&C-based explanations, examples, and graphs – pages 56-66
- Introduction to Credibility Theory, by Thomas N. Herzog, 1999
 - Not reviewed because it requires a purchase
- An Introduction to Credibility Theory, by L.H. Longley-Cook, 1962
 - This is a historical introduction to credibility from the early 1960s. While newer introductions may be more to the point, it is an interesting alternative.
 - A modification of the asymptotic formula is presented. The example shown calculates the credibility weight as $Z = 1.5N/(N + 500)$, where N is the number of claims. Calculates Z for $N = 1$ to 1,000; not applicable for $N > 1,000$.
- Loss Models, from Data to Decisions, Fourth Edition (2012), by Stuart Klugman, Harry Panjer, and Gordon Willmot. A fifth edition is in the works.
 - Likely a good foundation for accessing more advanced credibility methods
 - Starts with modeling, random variables, and basic quantities related to distributions, then explores three types of distributions: continuous, discrete, and advanced discrete
 - Adds insurance features such as frequency, severity, policy limits, coinsurance, and deductibles before covering aggregate loss models
 - Chapters then cover construction of empirical models and parametric statistical methods before exploring credibility and simulation
- Principles and Application of Credibility Theory, by Vincent Goulet, 1998
 - Mostly quite readable; formulas sometimes get very dense
 - Brief Historical Review – pages 9-13, interesting
 - Limited Fluctuation Credibility – pages 14-18, not too dense
 - Greatest Accuracy Credibility – pages 19-23, dense in places
 - Exact Bayesian Credibility – pages 24-30, mostly incomprehensible
 - Bühlmann-Straub Model – pages 31-40, includes a good example
 - Hierarchical Credibility – pages 41-48, good diagram on page 43
 - Crossed Classification Credibility – pages 49-58, impressive, difficult formulas
- Proposed Guidelines for Full Credibility, issued by CMS (Centers for Medicare and Medicaid Services), February 15, 2018
 - Proposes guidelines for full credibility in terms of the number of member months required for Medicare Advantage and Prescription Drugs. Defines full credibility as a 95% probability that the observed mean is within 10% of the true mean.

- Shows tables of number of member months required to achieve full credibility based on values of σ/μ (standard deviation divided by the mean) and average monthly exposure per member per year.

Appendix A: Calculation of $E[Var[X]]$ and $Var[E[X]]$

To illustrate the calculation of $E[Var[X]]$ and $Var[E[X]]$, we will examine a risk portfolio consisting of three kinds of dice: four-sided, six-sided, and eight-sided with, respectively, the numbers 1 through 4, 1 through 6, and 1 through 8 on the sides of the different kinds of dice. There are 100 dice in total, of which 60 are 4-sided, 30 are 6-sided, and 10 are 8-sided.

Please note that, unlike calculations based on experience data, this example is based purely on expected distributions.

A.1 Weights, Means, and Variance

To translate this into an insurance example, you might imagine a type of health or casualty insurance coverage that can pay multiple claims per year per insured. The 4-sided dice could then represent a risk group with insureds who will incur either 1, 2, 3, or 4 claims per year, with a 25% probability of each number of claims. X_1 will denote the random variable assigned to the first risk group, the 4-sided dice. We will assume that the elements of the X_1 vector take on one of the four values of 1, 2, 3, or 4 with equal probability.

The expected or mean number of claims per roll of the die can then be calculated as:

$$E[X_1] = \frac{1+2+3+4}{4} = 2.500.$$

The statistic $E[X_1^2]$ is needed to calculate variance; it is the average number squared per roll of the die:

$$E[X_1^2] = \frac{1^2 + 2^2 + 3^2 + 4^2}{4} = 7.500.$$

From the above two statistics, the variance, $Var[X_1]$, and standard deviation, σ_1 , can be calculated:

$$Var[X_1] = E[X_1^2] - (E[X_1])^2 = 7.500 - (2.500)^2 = 1.250.$$

$$\sigma_1 = \sqrt{Var[X_1]} = 1.118.$$

Each of the three risk groups has the following attributes and statistics:

- i identifies a risk group. For a mortality study, it might represent insureds covered by a particular insurance company. For the dice example, $i = 1$ will represent 4-sided dice, $i = 2$ will represent 6-side dice, and $i = 3$ will represent 8-sided dice.
- N_i is the exposure for risk group i . For a mortality study, it would equal the number of lives exposed. In the dice example, N_i will take on values of 60, 30, and 10 for $i = 1, 2$, and 3, respectively.
- N , the portfolio's number of observations, is equal to the sum of N_i for all i .
- Wt_i , the weight for risk group i , is calculated as N_i/N .

- X_i , the risk group's random variable under study, is a vector containing the risk group's observations.
 - For a mortality study, each element in the vector would contain a death count of 0 or 1 for each of the N_i lives in the risk group.
 - In the dice example, the random variable will contain results from repeated rolls of the die, such as the numbers 1 through 4 for the 4-sided die or 1 through 8 for the 8-sided die.
- $E[X_i]$, the expectation of X_i , is the mean or average value of the observations for the risk group.
 - For a mortality study, this would be the observed mortality rate.
 - In the dice example, the four-sided die has a true mean of 2.5, but actual results will vary: When you randomly roll a four-sided die 60 times, it is very unlikely that the average will exactly equal 2.5.
- $E[X_i^2]$, the expectation of X_i^2 , is the mean of the square of each of the risk group's observations. It is calculated by squaring each observed value, summing up those squares and then dividing by N_i .
 - For a random variable that takes on values of only 0 and 1, such as a death count, $E[X_i^2]$ is equal to $E[X_i]$, since $0^2 = 0$ and $1^2 = 1$.
- $Var[X_i]$, the variance of X_i , is calculated as $E[X_i^2] - (E[X_i])^2$.
- σ_i , the standard deviation of X_i is calculated as $\sqrt{Var[X_i]}$.

To more simply illustrate the calculations, the following table shows risk group statistics based on the true means, $E[X_i]$ and $E[X_i^2]$, rather than observed data:

RISK GROUP WEIGHTS WITH TRUE MEANS AND VARIANCES							
Risk Group (i)	Sides per Die	N_i Number of Rolls	Wt _i Risk Group Weight	$E[X_i]$ Risk Group Mean	$E[X_i^2]$ Risk Group Mean of X^2	$Var[X_i]$ Risk Group Variance	σ_i Risk Group Std Dev
1	4	60	60%	2.5000	7.5000	1.2500	1.1180
2	6	30	30%	3.5000	15.1667	2.9167	1.7078
3	8	10	10%	4.5000	25.5000	5.2500	2.2913
Total		100	100%				

A.2 Portfolio Statistics

The next step is to calculate statistics for the portfolio as a whole:

- X , the portfolio's random variable under study, is composed of the risk group random variable vectors X_i of length N_i . For a mortality study, each element in this collection of vectors would contain a death count of 0 or 1 for each of the N lives in the portfolio. For the dice example, X would contain the results of 100 rolls from the three types of dice.
- $E[X]$, the expectation or mean of X , is calculated by giving equal weight to each observation. It can also be calculated as the weighted average of the risk group means.
- $E[X^2]$, the expectation or mean of X^2 , is calculated by giving equal weight to each observation squared. It can also be calculated as the weighted average of the risk group expectations of X_i^2 , $E[X_i^2]$.
- $E[(E[X])^2]$, the expectation of risk group means squared, is calculated as the weighted average of the squares of the risk group means, $(E[X_i])^2$.

The numerator of the constant K is v , *the expected value of the process variance*. It is equal to the expectation of the variance of X :

$$\begin{aligned} v = E[Var[X]] &= E[E[X^2] - (E[X])^2] = E[E[X^2]] - E[(E[X])^2] \\ &= E[X^2] - E[(E[X])^2]. \end{aligned}$$

In contrast, the variance of X is calculated as:

$$Var[X] = E[X^2] - (E[X])^2.$$

The difference between $E[Var[X]]$ and $Var[X]$ is the difference between the expectation of $(E[X])^2$, denoted as $E[(E[X])^2]$, and $(E[X])^2$.

The following table shows the portfolio's statistics based on true means (both $E[X_i]$ and $E[X_i^2]$) for each risk group, rather than observed data:

PORTFOLIO STATISTICS BASED ON TRUE MEANS					
Risk Group (i)	$E[E[X]] = E[X]$ Wtd Mean	$E[E[X^2]] = E[X^2]$ Wtd Mean of X^2	$E[(E[X])^2]$ $Wtd (Mean)^2$	$E[Var[X]] = Wtd \text{ Mean of } X^2 - Wtd (Mean)^2$	$Var[X] = E[X^2] - (E[X])^2$
1	1.5000	4.5000	3.7500	0.7500	
2	1.0500	4.5500	3.6750	0.8750	
3	0.4500	2.5500	2.0250	0.5250	
Total	3.0000	11.6000	9.4500	2.1500	2.6000

Appendix B: Minimizing Mean-Squared Error

The following formulas are a simplified version of those developed on pages I.11 to I.15 of Credibility Theory Practices, published by the Society of Actuaries in December 2009.

We will use the following variables for a portfolio consisting of risk groups:

- μ_r is the true, but unknown, mean of risk group r ,
- \hat{u}_r is the observed mean of risk group r ,
- Z is the credibility weight, and
- W is a constant to be solved for.

Additionally:

- μ will be used to signify $E[\mu_r]$, the expectation of μ_r and
- σ^2 will be used to signify $Var[\mu_r]$, the variance of μ_r .

$E[\hat{u}_r]$, the expectation of \hat{u}_r , is therefore equal to μ , because its true mean is μ_r .

We will assume that the best estimate for the r^{th} risk group's *Credibility-Weighted Rate* will take the following linear form:

$$\text{Credibility-Weighted Rate} = Z\hat{u}_r + W.$$

The difference or error between the true mean and the best estimate is:

$$\text{Error} = \mu_r - Z\hat{u}_r - W$$

Z and W will be selected to minimize the expected mean-squared error, equal to the following:

$$E[(\mu_r - Z\hat{u}_r - W)^2]$$

The minimum error will be assumed to occur where the partial derivatives with respect to Z and W are zero. Setting the partial derivatives equal to zero provides two linear equations with two unknowns, allowing Z and W to be solved for.

Taking the partial derivative of $E[(\mu_r - Z\hat{u}_r - W)^2]$ with respect to W and setting the result equal to zero yields the following formula:

$$-2E[\mu_r - Z\hat{u}_r - W] = 0.$$

Rearranging terms, we have:

$$W = E[\mu_r] - ZE[\hat{u}_r].$$

Substituting μ for $E[\mu_r]$ and $E[\hat{u}_r]$, we have:

$$W = (1 - Z)\mu.$$

Combining the above result with the formula for *Credibility-Weighted Rate* presented earlier in this section, we have:

$$\text{Credibility-Weighted Rate} = Z\hat{u}_r + (1 - Z)\mu.$$

The above formula closely matches the formula for the *Credibility-Weighted Rate* presented in Section 3.2:

$$\text{Credibility-Weighted Rate} = (Z)\text{Observed Rate} + (1 - Z)\text{Portfolio Rate}.$$

By definition, \hat{u}_r is the *Observed Rate*, so the implication is that μ and *Portfolio Rate* are one and the same. This would indicate that the *Portfolio Rate* is assumed to be the best estimate of the underlying true rate for each of the risk groups.

Next, we take the partial derivative of $E[(\mu_r - Z\hat{u}_r - W)^2]$ with respect to Z and set the result equal to zero to yield the following formula:

$$-2E[(\mu_r - Z\hat{u}_r - W)\hat{u}_r] = 0.$$

Rearranging terms, we have:

$$W = \frac{E[\mu_r\hat{u}_r] - ZE[\hat{u}_r^2]}{E[\hat{u}_r]}$$

$E[\mu_r\hat{u}_r]$ can be shown to be equal to $E[\mu_r^2]$. Because $\text{Var}[X] = E[X^2] - (E[X])^2$, we know that $\sigma^2 = E[\mu_r^2] - \mu^2$. Making these two substitutions into the prior equation and replacing $E[\hat{u}_r]$ with μ , we have:

$$W = \frac{\mu^2 + \sigma^2 - ZE[\hat{u}_r^2]}{\mu}$$

There is no simple substitution for $E[(\hat{u}_r)^2]$: It is typically the result of a complex calculation. Combining the above formula with the prior result, $W = (1 - Z)\mu$, we can eliminate W :

$$(1 - Z)\mu = \frac{\mu^2 + \sigma^2 - ZE[\hat{u}_r^2]}{\mu}$$

Solving for Z , we have:

$$Z = \frac{\sigma^2}{E[\hat{u}_r^2] - \mu^2}$$

Because $\text{Var}[X] = E[X^2] - (E[X])^2$, we know that $\sigma^2 = E[\mu_r^2] - \mu^2$. Substituting this relationship into the prior equation, we have:

$$Z = \frac{\sigma^2}{\sigma^2 + E[\hat{u}_r^2] - E[\mu_r^2]}.$$

Notice how the above equation resembles the Asymptotic formula, $Z = N/(N + K)$. By multiplying the numerator and denominator by the ratio N/σ^2 , we can isolate the formula for K as everything that follows N in the denominator:

$$Z = \frac{N}{N + \frac{N}{\sigma^2} (E[\hat{u}_r^2] - E[\mu_r^2])}.$$

We therefore have:

$$K = \frac{N}{\sigma^2} (E[\hat{u}_r^2] - E[\mu_r^2]).$$

About the Society of Actuaries

The Society of Actuaries (SOA), formed in 1949, is one of the largest actuarial professional organizations in the world dedicated to serving 30,000 actuarial members and the public in the United States, Canada, and worldwide. In line with the SOA Vision Statement, actuaries act as business leaders who develop and use mathematical models to measure and manage risk in support of financial security for individuals, organizations, and the public.

The SOA supports actuaries and advances knowledge through research and education. As part of its work, the SOA seeks to inform public policy development and public understanding through research. The SOA aspires to be a trusted source of objective, data-driven research and analysis with an actuarial perspective for its members, industry, policymakers, and the public. This distinct perspective comes from the SOA as an association of actuaries, who have a rigorous formal education and direct experience as practitioners as they perform applied research. The SOA also welcomes the opportunity to partner with other organizations in our work where appropriate.

The SOA has a history of working with public policymakers and regulators in developing historical experience studies and projection techniques as well as individual reports on health care, retirement, and other topics. The SOA's research is intended to aid the work of policymakers and regulators and follow certain core principles:

Objectivity: The SOA's research informs and provides analysis that can be relied upon by other individuals or organizations involved in public policy discussions. The SOA does not take advocacy positions or lobby specific policy proposals.

Quality: The SOA aspires to the highest ethical and quality standards in all of its research and analysis. Our research process is overseen by experienced actuaries and non-actuaries from a range of industry sectors and organizations. A rigorous peer-review process ensures the quality and integrity of our work.

Relevance: The SOA provides timely research on public policy issues. Our research advances actuarial knowledge while providing critical insights on key policy issues and thereby provides value to stakeholders and decision makers.

Quantification: The SOA leverages the diverse skill sets of actuaries to provide research and findings that are driven by the best available data and methods. Actuaries use detailed modeling to analyze financial risk and provide distinct insight and quantification. Further, actuarial standards require transparency and the disclosure of the assumptions and analytic approach underlying the work.