Credibility Theory

Credibility theory (*see* **Insurance Pricing/Nonlife**) refers to a body of techniques actuaries use to assign premiums to individual policyholders in a heterogeneous portfolio. These techniques are applicable in fields other than insurance wherever one needs to distinguish between individual and group effects in a set of nonidentically distributed data. As such, credibility theory is sometimes defined shortly as "the mathematics of heterogeneity".

With origins dating back to 1914, credibility theory is widely recognized as the cornerstone of casualty actuarial science (*see* **Dependent Insurance Risks**; **Reinsurance**; **Securitization/Life**). It is divided in two main approaches, each representing a different manner of incorporating individual experience in the ratemaking process (*see* **Nonlife Loss Reserving**). The oldest approach is limited fluctuations credibility, where the premium of a policyholder is based solely on its own experience, provided its experience is stable enough to be considered fully credible. The theory behind limited fluctuations credibility is rather simple and the technique has few correct applications, so it will be reviewed only briefly in this article.

The second approach concentrates on the homogeneity of the portfolio rather than the stability of the experience to determine the best premium for a policyholder. In greatest accuracy credibility, individual experience is taken into consideration only if it is significantly different from that of the portfolio. The more heterogeneous the portfolio, the more important becomes individual experience, and *vice versa*.

Experience Rating

The first concern of an insurer – whether a public or a private corporation – when building a tariff is to charge enough premiums to fulfill its obligations. For various reasons, including competitiveness, the insurer may then seek to distribute premiums fairly between policyholders. A classification structure (for example: according to age, sex, type of car, etc.) is usually the first stage in premium distribution. Experience-rating (*see* **Risk Classification in Nonlife Insurance**) systems, in general, and credibility methods, in particular, then constitute an efficient second stage.

As the name suggests, an experience-rating system takes the individual experience of a policyholder into account to determine its premium. Such systems require the accumulation of a significant volume of experience. Experience rating is thus especially used in workers compensation and automobile insurance.

On a more formal basis, Bühlmann [1] proposes the following definition:

> Experience rating aims at assigning to each individual risk its own *correct premium* (rate). The *correct premium* for any period depends exclusively on the (unknown) claims distribution of the individual risk for this same period.

Example 1 Consider this simplified example (more details are presented in [2] or [3]): a portfolio is composed of ten policyholders and each policyholder can incur at most one claim of amount 1 per year. Before observing any experience for this portfolio, the insurer considers the policyholders equivalent on a risk level basis. Expecting that, on average, the portfolio will incur two claims per year, the insurer charges an identical premium of 0.20 to each policyholder. This is the collective premium.

The experience for this portfolio after 10 years is shown in Table 1. The insurer observed a total of 23 claims, for an average cost per policyholder of $23/100 = 0.23$. However, these 23 claims are not uniformly distributed among policyholders: policyholders 7, 8, and 10 incurred no claims, while policyholder 9 alone incurred seven.

Table 1 Experience of the simplified portfolio of Example 1 after 10 years

	Policyholder									
Year	1	2	3	4	5	6	7	8	9	10
1									1	
2	1	1	1						1	
3	1								1	
4			1						1	
5									1	
6		1								
7	1	1		1	1					
8	1			1		1			1	
9	1				1					
10	1								1	
$\bar{S}_i$	0.6	0.3	0.2	0.2	0.2	0.1	0	0	0.7	0
$\bar{S}$					0.23					

Therefore, the collective premium in this example is globally adequate, but clearly not fair. To avoid antiselection, the insurer should charge more to some policyholders and less to others. Experience showed that the portfolio is, to some degree, heterogeneous.

One could very well restate this example in a context other than insurance. The data of Table 1 merely establish whether an "event" occurred or not in 10 "trials" and for 10 different "subjects". The goal is then to estimate the probability of occurrence for trial 11 for each subject, knowing that the subjects are not "equivalent".

There exist numerous experience-rating systems, including bonus–malus and merit–demerit systems, participating policies or commissions in reinsurance [1, 4, 5], but the most widely used methods are credibility models.

Limited Fluctuations Credibility

Limited fluctuations credibility originated in the early 1900s with Mowbray's paper "How Extensive A Payroll Exposure Is Necessary To Give A Dependable Pure Premium?" As the title suggests, Mowbray was interested in finding a level of payroll in workers compensation insurance for which the pure premium of a given employer would be considered fully dependable or, in other words, fully credible.

We may believe that this question arose from one or a few large employers requesting from their insurer to pay a premium better tailored to their own experience rather than to the experience of their group. The employers would rightfully argue that given their large size and by virtue of the law of large numbers their experience is fairly stable in time, thus allowing the insurer to compute more accurate and fair pure premiums.

Mowbray [6] defines a dependable pure premium to be "one for which the probability is high that it does not differ from the [true] pure premium by more than an arbitrary limit". Translated in modern mathematical terms, this is essentially asking that

$$\Pr[(1-k)E[S] \leq S \leq (1+k)E[S]] \geq p \tag{1}$$

for a small value of k and a large probability p. Here, the random variable S represents the experience of a policyholder, in one form or another. We detail three typical examples below.

1. Mowbray defined S as the number of accidents of an employer, assuming a binomial distribution with parameters n, the number of employees, and θ, the known probability of accident. Solving equation (1) for n, using the central limit theorem yields

$$n \geq \left(\frac{\zeta_{(1-p)/2}}{k}\right)^2 \frac{1-\theta}{\theta} \tag{2}$$

where ζ_α is the $100(1-\alpha)$th percentile of the standard normal distribution.

2. The size of a policyholder may be defined as the expected number of claims in a given period (typically, 1 year). For such cases, one usually defines S as the total amount of claims in the period and further assumes that the distribution of S is a compound Poisson. Solving equation (1) for the Poisson parameter λ yields

$$\lambda \geq \left(\frac{\zeta_{(1-p)/2}}{k}\right)^2 \left(1 + \frac{\operatorname{var}[X]}{E[X]^2}\right) \tag{3}$$

where X is the random variable of claim amounts.

3. The two examples above determine a policyholder's admissibility to full credibility one period at a time. An alternative criterion is the number of periods of experience. This requires to set S in equation (1) as the average total amount of claims after n years:

$$S = \frac{S_1 + S_2 + \cdots + S_n}{n} \tag{4}$$

where $S_1, \ldots, S_n$ are independent and identically distributed random variables of the total amount of claims per period. Then the experience of a policyholder is considered fully credible after

$$n \geq \left(\frac{\zeta_{(1-p)/2}}{k}\right)^2 \frac{\operatorname{var}[S_i]}{E[S_i]^2} \tag{5}$$

periods of experience.

The interested reader will find more examples in [7].

A policyholder's experience is thus considered fully credible if it fluctuates moderately from one period to another. That is, the credibility criterion is stability. This stability of the experience usually increases with the volume of the policyholder, whether it is expressed in premium volume, number of claims, number of employees, number of years of experience, or any other exposition base.

Truly appropriate applications of Mowbray's procedure are rare. In insurance, it should be reserved for applications where the stability of the experience is of foremost importance. One good example is the determination of an admissibility threshold to a retrospective insurance system, where the policyholder's premium is readjusted at the end of the year after the total claim amount is known.

To this day, limited fluctuations remains the credibility procedure most widely used by American actuaries. It has been extended in various and often *ad hoc* ways. One such extension deals with the case of a policyholder only partially reaching the full credibility criterion. As proposed originally by Whitney in his seminal paper of 1918 [8], one charges a pure premium π that is a weighted average of the individual experience S and the collective premium m:

$$\pi = zS + (1 - z)m \tag{6}$$

Here, $0 \le z \le 1$ is the so-called credibility factor.

Over the years, many partial credibility formulas have been proposed. Among the most widely used are the following:

$$z = \min\left\{\sqrt{\frac{n}{n_0}}, 1\right\} \tag{7}$$

$$z = \min\left\{\left(\frac{n}{n_0}\right)^{2/3}, 1\right\} \tag{8}$$

and

$$z = \frac{n}{n + K} \tag{9}$$

where n_0 is the full credibility level and K is a constant determined upon judgment, usually to limit the size of the premium change from one year to the next. The third formula is the one introduced by Whitney and the only one in which the (partial) credibility level never reaches unity.

It should be emphasized that in such a context, partial credibility does not seek to find the most accurate premium for a policyholder. The goal is rather to incorporate in the premium as much individual experience as possible while still keeping the premium sufficiently stable. When credibility is used to find the best estimate of a policyholder's pure risk premium, one should turn toward greatest accuracy methods.

Greatest Accuracy Credibility

The greatest accuracy approach to credibility theory seeks to find the "best" (in a sense yet to be defined) premium to charge a policyholder. This is achieved by distributing the collective premium in an optimal way among the members of a group.

It is now well recognized that the first paper on greatest accuracy credibility was the aforementioned 1918 paper by Whitney [8]. Whitney developed formula (6) out of "the necessity, from the standpoint of equity to the individual risk, of striking a balance between class-experience on the one hand and risk-experience on the other." How this balance is calculated does not depend solely on the stability of the experience, like in the limited fluctuations approach, but rather on the homogeneity of the portfolio. Indeed, Whitney writes:

> There would be no experience-rating problem if every risk within the class were typical of the class, for in that case the diversity in the experience would be purely adventitious.

However, Whitney was ahead of his time and his ideas were either not well received or not well understood. For example, he was criticized for using an early version of the Bayes rule; see [9]. For decades, actuaries essentially remembered and used formula (6) and the form $z = n/(n + K)$ for the credibility factor.

The greatest accuracy approach lay dormant until the publication of Arthur L. Bailey's papers [10, 11]. The second paper is an especially enlightening exposition of the state of statistical and actuarial practices at the time. In particular, with respect to credibility theory, Bailey writes:

> The trained statistician cries "Absurd! Directly contrary to any of the accepted theories of statistical estimation." The actuaries themselves have to admit that they have gone beyond anything that has been proven mathematically, that all of the values involved are still selected on the basis of judgment, and that the only demonstration they can make is that, in actual practice, it works.

Bailey then poses himself as an advocate of the then controversial Bayesian philosophy (*see* **Natural**

Resource Management; **Risk in Credit Granting and Lending Decisions: Credit Scoring**; **Reliability Demonstration**; **Cross-Species Extrapolation**), arguing that the notion of prior opinion is natural in actuarial work:

> At present, practically all methods of statistical estimation appearing in textbooks on statistical methods or taught in American universities are based on an equivalent to the assumption that any and all collateral information or *a priori* knowledge is worthless. ... Philosophers have recently discussed the credibilities to be given to various elements of knowledge [12], thus undermining the accepted philosophy of the statisticians. However, it appears to be only in the actuarial field that there has been an organized revolt against discarding all prior knowledge when an estimate is to be made using newly acquired data.

Bailey then shows that the Bayesian estimator of the policyholder's true risk premium obtained by minimizing the mean square error is a credibility premium of the form of equation (6) for certain combinations of distributions. The credibility factor is still of the form $z = n/(n + K)$, where K depends on the parameters of the model.

Modeling Heterogeneity

The classical mathematical model of greatest accuracy credibility was formalized by Bühlmann [1, 5]. Consider a heterogeneous portfolio (group) of I policyholders. The risk level of policyholder $i = 1, \ldots, I$ (whether it is a good or a bad driver, for example) is unknown, but past claims data $S_{i1}, \ldots, S_{in}$ are available for ratemaking purposes.

We make the following assumptions:

1. Claim amounts of policyholder i are (conditionally) independent and identically distributed with cumulative distribution function (cdf) $F(x|\theta_i)$. Parameter θ_i is the realization of a random variable Θ_i.
2. Random variables $\Theta_1, \ldots, \Theta_I$ are identically distributed with cdf $U(\theta)$.
3. The policyholders are independent, meaning that the claims record of one policyholder has no impact on the claims record of another.

The random variable Θ_i represents the risk level of policyholder i. This is an abstract and nonobservable random variable – otherwise, the ratemaking problem would be easily solved. The portfolio is heterogeneous since each policyholder has its own risk level. Yet, the identical distribution assumption for the random variables $\Theta_1, \ldots, \Theta_I$ means that the risk levels all come from the same process. Hence the policyholders are similar enough to justify grouping them in the same portfolio.

In a pure Bayes setting, $U(\theta)$ represents the prior distribution of the risk levels. The distribution is revised – yielding the posterior distribution – as new claims data become available. In the more practical empirical Bayes setting, $U(\theta)$ is rather seen as the structure function of the portfolio, that is the distribution of risk levels within the group of policyholders.

Prediction

The goal in greatest accuracy credibility is to compute the best prediction of the future claims $S_{i,n+1}$ for every policyholder. If the risk level of policyholder i were known, the best (in the mean square sense) prediction would be the expected value

$$\mu(\theta_i) = E[S_{it}|\Theta_i = \theta_i] = \int_0^\infty x \, \mathrm{d}F(x|\theta_i) \qquad (10)$$

In the actuarial literature, this function is called the *risk premium* (*see* **Equity-Linked Life Insurance**; **Inequalities in Risk Theory**; **Risk Attitude**). Now, the risk levels and, consequently, the risk premiums are unknown. One is thus left with the equivalent problems of predicting future claims $S_{i,n+1}$ or finding approximations of the risk premiums $\mu(\theta_i)$.

A first approximation of the risk premiums is the weighted average of all possible risk premiums:

$$m = E[\mu(\Theta)] = \int_{-\infty}^\infty \mu(\theta) \, \mathrm{d}U(\theta) \qquad (11)$$

This approximation will be the same for all policyholders. It is the *collective premium* (*see* **Premium Calculation and Insurance Pricing**; **Ruin Probabilities: Computational Aspects**).

As explained earlier, the collective premium, although globally adequate, fails to achieve an optimal premium distribution among policyholders. In statistical terms, this means that there exist better approximations of the risk premiums when experience is available. Indeed, the best approximation (or estimation, or prediction) of the risk premium $\mu(\theta_i)$ is the function $g^*(S_{i1}, \ldots, S_{in})$ minimizing the mean square error

$$E[(\mu(\Theta) - g(S_{i1}, \ldots, S_{in}))^2] \tag{12}$$

where $g(\cdot)$ is any function. Most standard mathematical statistics text (see, e.g. [13]) show that function $g^*(S_{i1}, \ldots, S_{in})$ is the so-called Bayesian premium

$$\begin{aligned} B_{i,n+1} &= E[\mu(\Theta)|S_{i1}, \ldots, S_{in}] \\ &= \int_{-\infty}^{\infty} \mu(\theta)\, \mathrm{d}U(\theta|S_{i1}, \ldots, S_{in}) \end{aligned} \tag{13}$$

Function $U(\theta|x_1, \ldots, x_n)$ is the aforementioned posterior distribution of the risk levels, or the revised structure function of the portfolio, after claims experience became available.

Let $f(x|\theta) = F'(x|\theta)$ be the conditional probability density function (pdf) or probability mass function (pmf) of the claim amounts and $u(\theta) = U'(\theta)$ be the pdf or pmf of the risk levels. Then, the posterior distribution of the risk levels is obtained from the prior distribution using the Bayes rule:

$$u(\theta_i|x_1, \ldots, x_n) = \frac{f(x_1, \ldots, x_n|\theta_i)u(\theta_i)}{\displaystyle\int_{-\infty}^{\infty} f(x_1, \ldots, x_n|\theta)\, \mathrm{d}U(\theta)} \tag{14}$$

By conditional independence of claim amounts, this can be rewritten as

$$u(\theta_i|x_1, \ldots, x_n) = \frac{\displaystyle\prod_{t=1}^{n} f(x_t|\theta_i)u(\theta_i)}{\displaystyle\int_{-\infty}^{\infty} \prod_{t=1}^{n} f(x_t|\theta)\, \mathrm{d}U(\theta)} \tag{15}$$

The Bayesian premium $B_{i,n+1}$ is thus the best computable prediction of the future claims $S_{i,n+1}$. Akin to the collective premium, it is a weighted average of all possible risk premiums, but using the posterior rather than the prior distribution of the risk levels as weighting function. Reversing the argument, one can also see the collective premium as the Bayesian premium of the first year, when no experience is available.

Example 2 Consider policyholder 1 of the simplified portfolio of Example 1. This policyholder can incur at most one claim of amount 1 per period, but with unknown probability. If S_t is the random variable of the experience in year $t = 1, \ldots, n$, then $S_t|\Theta = \theta$ has a Bernoulli distribution with parameter θ. Parameter θ is seen as an outcome of a random variable Θ having a beta prior distribution with parameters α and β. That is, we have

$$f(x_t|\theta) = \theta^x(1-\theta)^{1-x}, \quad x = 0, 1 \tag{16}$$

and

$$u(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, \theta^{\alpha-1}(1-\theta)^{\beta-1}, \quad 0 < \theta < 1 \tag{17}$$

The risk premium, or average claim amount, is $\mu(\theta) = E[S_t|\Theta = \theta] = \theta$. This information is of no real use since the true value of θ is unknown. One can however estimate the average claim amount by the collective premium

$$m = E[\mu(\Theta)] = E[\Theta] = \frac{\alpha}{\alpha+\beta} \tag{18}$$

This is the best approximation in the first year, when no other information is available. However, the results of Table 1 show that the collective premium is inappropriate for this policyholder.

The accumulation of experience allows one to better assess the true value of θ by means of equation (15). The posterior distribution of Θ after n years of experience is, up to a proportionality constant,

$$\begin{aligned} u(\theta|x_1, \ldots, x_n) &\propto \theta^{\alpha-1}(1-\theta)^{\beta-1} \\ &\quad \times \prod_{t=1}^{n} \theta^{x_t}(1-\theta)^{1-x_t} \qquad (19) \\ &= \theta^{\alpha+\sum_{t=1}^{n} x_t - 1}(1-\theta)^{\beta+n-\sum_{t=1}^{n} x_t - 1} \qquad (20) \end{aligned}$$

That is, the distribution of $\Theta|S_1 = x_1, \ldots, S_n = x_n$ is a beta distribution with updated parameters $\tilde{\alpha} = \alpha + \sum_{t=1}^{n} x_t$ and $\tilde{\beta} = \beta + n - \sum_{t=1}^{n} x_t$. Therefore, the Bayesian premium for year $n+1$ is

$$B_{n+1} = E[\mu(\Theta)|S_1, \ldots, S_n] \tag{21}$$

$$= E[\Theta | S_1, \ldots, S_n] \tag{22}$$

$$= \frac{\tilde{\alpha}}{\tilde{\alpha} + \tilde{\beta}} \tag{23}$$

$$= \frac{\alpha + \sum_{t=1}^{n} S_t}{\alpha + \beta + n} \tag{24}$$

Setting $\alpha = 1$ and $\beta = 4$ in the above results gives a collective premium of 0.2, as in Example 1. Figure 1 shows the evolution of the probability of claim distribution from the prior to the posterior after 10 years of experience. One sees that the distribution gets increasingly concentrated around the true value of θ. Accordingly, the Bayesian premiums in Table 2 are getting closer to the true risk level of the policyholder.

Table 2 Bayesian premiums for policyholder 1 of Table 1 using a beta prior distribution with $\alpha = 1$ and $\beta = 4$

n	x_n	$\sum_{t=1}^{n} x_t$	$\alpha + \sum_{t=1}^{n} x_t$	$\beta + n - \sum_{t=1}^{n} x_t$	B_{n+1}
0	–	–	1	4	0.200
1	0	0	1	5	0.167
2	1	1	2	5	0.286
3	1	2	3	5	0.375
4	0	2	3	6	0.333
5	0	2	3	7	0.300
6	0	2	3	8	0.273
7	1	3	4	8	0.333
8	1	4	5	8	0.385
9	1	5	6	8	0.429
10	1	6	7	8	0.467

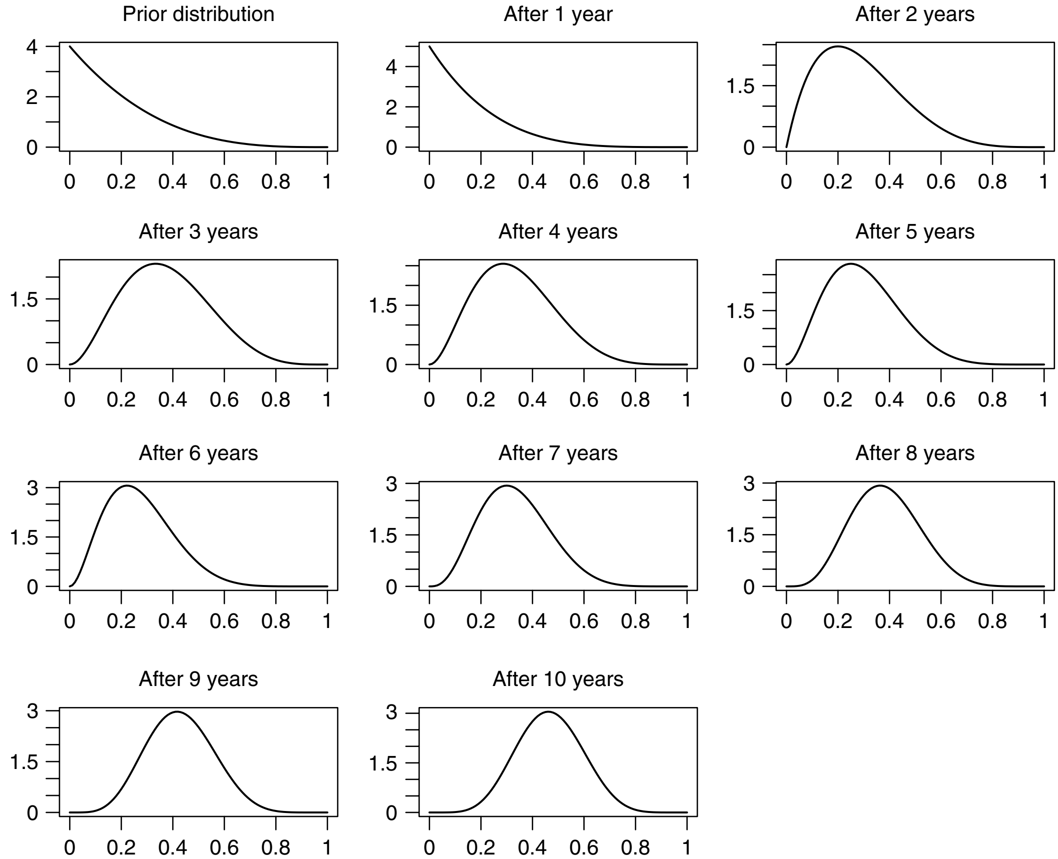

Figure 1 Posterior distributions for policyholder 1 of Table 1 using a beta prior distribution with $\alpha = 1$ and $\beta = 4$

Linear Bayes Credibility

The Bayesian premium equation (24) can be rewritten as

$$B_{n+1} = \frac{n}{n+\alpha+\beta}\bar{S} + \frac{\alpha+\beta}{n+\alpha+\beta}\frac{\alpha}{\alpha+\beta} \quad (25)$$

$$= z\bar{S} + (1-z)m \quad (26)$$

with $z = n/(n+\alpha+\beta)$, the same form as equation (6). A premium of this form is called a *credibility premium* and z is the *credibility factor*.

Whitney [8] and Bailey [11] were the first to show that the Bayesian premium is a credibility premium for certain combinations of distributions and, furthermore, that the credibility factor is always of the form $z = n/(n+K)$, with K some constant. Bailey's partial results were extended by Mayerson [14] and later unified by Jewell [15].

The five main combinations of distributions giving a linear Bayesian premium are given in Table 3 (see [3] for a more complete table). Note that in all cases the posterior distribution of the risk parameter is of the same form as the prior, but with revised parameters.

The credibility premium is a weighted average between the collective premium and the individual average of a policyholder. The weight given to the latter increases with the number of years of experience (n). This is both simple and intuitively sound, two desirable properties of the credibility premium.

The Bühlmann and Bühlmann–Straub Models

In practice, the Bayesian premium equation (13) is of limited interest for two main reasons. First, one has to make assumptions on the distributions of $S_{it}|\Theta_i$ and Θ_i, something that can be difficult and highly subjective. Second, the Bayesian premium can be complicated to obtain – see [2] for what happens when the beta distribution in Example 2 is replaced by an apparently simple uniform distribution on (a, b) – and does not in general lie between the individual and collective premiums.

Bühlmann [5] proposed to restrict the approximation of the risk premium to linear functions of the observations. The optimal linear prediction happens to be a credibility premium

$$\pi^B_{i,n+1} = z\bar{S}_i + (1-z)m \quad (27)$$

with

$$z = \frac{n}{n+s^2/a} \quad (28)$$

$$s^2 = E[\sigma^2(\Theta)] \quad (29)$$

$$a = \mathrm{var}[\mu(\Theta)] \quad (30)$$

and $\sigma^2(\Theta_i) = \mathrm{var}[S_{it}|\Theta_i]$.

The so-called structure parameters m, s^2 and a define the inner structure of the portfolio:

1. m measures the global average of the portfolio;
2. s^2 measures the variability within policyholders, that is of claims in time;

Table 3 Combinations of distributions yielding a linear Bayesian premium and main results

$f(x \mid \theta)$	$u(\theta)$	B_{n+1}	z
Bernoulli(θ)	Beta(α, β)	$\frac{\alpha+\sum_{t=1}^n S_t}{\alpha+\beta+n}$	$\frac{n}{n+\alpha+\beta}$
Geometric(θ)	Beta(α, β)	$\frac{\beta+\sum_{t=1}^n S_t}{\alpha+n-1}$	$\frac{n}{n+\alpha-1}$
Poisson(θ)	Gamma(α, λ)	$\frac{\alpha+\sum_{t=1}^n S_t}{\lambda+n}$	$\frac{n}{n+\lambda}$
Exponential(θ)	Gamma(α, λ)	$\frac{\lambda+\sum_{t=1}^n S_t}{\alpha+n-1}$	$\frac{n}{n+\alpha-1}$
Normal(θ, σ_2^2)	Normal(μ, σ_1^2)	$\frac{\sigma_1^2\sum_{t=1}^n S_t+\sigma_2^2\mu}{n\sigma_1^2+\sigma_2^2}$	$\frac{n}{n+\sigma_2^2/\sigma_1^2}$

3. a measures the variability between policyholders, that is the heterogeneity of the portfolio.

One can see that more weight will be given to individual experience in the following scenarios:

1. the number of periods of experience (n) is large. This is justified since the experience of a policyholder represents its true risk level in the long run;
2. the value of parameter s^2 is small, the claims experience is globally stable in time. The individual averages $\bar{S}_i$ are thereby indicative of the risk levels, hence reducing the need for the collective premium;
3. the value of a is large, the portfolio is heterogeneous. Therefore, the individual averages are more accurate approximations of the risk premiums than the collective premium.

Figures 2 and 3 provide a graphical interpretation of the last two points. Each curve in these figures represents the experience of one policyholder. In each case, the credibility factor is largest in the figures marked (b).

The Bühlmann–Straub model (*see* **Insurance Pricing/Nonlife**) [16] is a generalization of the Bühlmann model allowing for variable exposure. This is essential in lines of business in which the size of the policyholders can vary within a portfolio. Consider for example workers compensation insurance where policyholders are employers: the exposure to risk of an employer with 1000 employees is much larger than that of an employer with only 10 employees.

In the Bühlmann–Straub model, a weight w_{it} (the number of employees or the payroll, for example) is assigned to each observation that we will now denote by X_{it}. One would expect that the experience of the larger policyholder will be more stable in time. To reflect this, the assumption of identical distribution used in Bühlmann's model is relaxed to let the conditional variance of an observation be inversely proportional to its weight:

$$\operatorname{var}[X_{it}|\Theta_i = \theta_i] = \frac{\sigma^2(\theta_i)}{w_{it}} \tag{31}$$

For this relation to hold, the observations X_{it} must now be ratios. Typically, the data will be defined as claim amounts divided by the exposure base:

$$X_{it} = \frac{S_{it}}{w_{it}} \tag{32}$$

The credibility premium differs from equation (27) in two respects only: the individual premium is estimated by a weighted average, and the credibility factor is a function of the total weight of a policyholder. We have

$$\pi_{i,n+1}^{\mathrm{BS}} = z_i X_{iw} + (1 - z)m \tag{33}$$

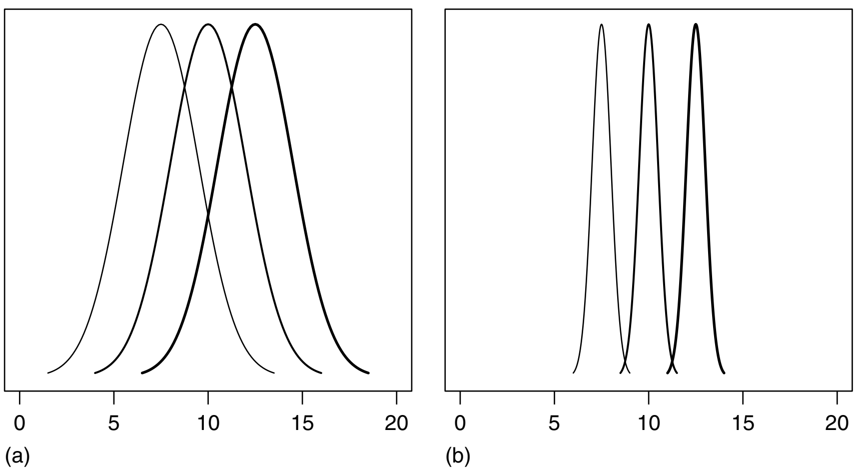

Figure 2 Effect of $s^2 = E[\sigma^2(\Theta)]$ on the credibility factor. (a) Large s^2, experience is too volatile to be reliable. (b) Small s^2, individual averages are more reliable. Means are the same in (a) and (b)

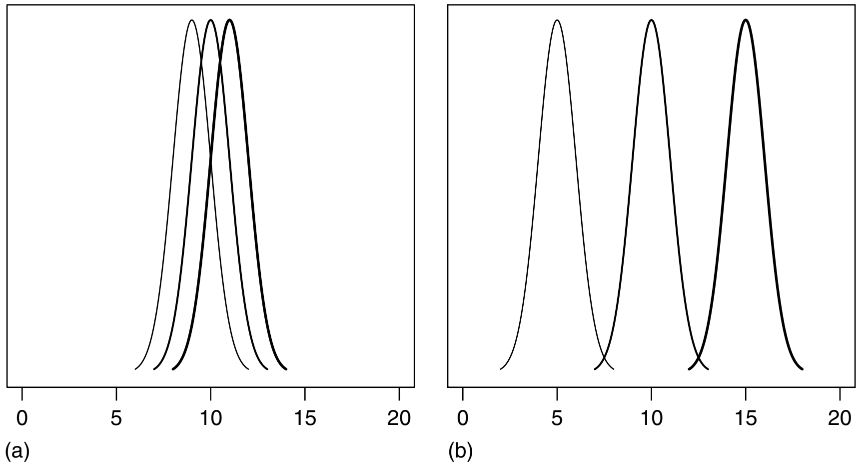

Figure 3 Effect of $a = \text{var}[\mu(\Theta)]$ on the credibility factor. (a) Small a, the portfolio is homogeneous. (b) Large a, the portfolio is heterogeneous. Variances are the same in (a) and (b)

with

$$z_i = \frac{w_{i\Sigma}}{w_{i\Sigma} + s^2/a} \tag{34}$$

$$X_{iw} = \sum_{t=1}^{n} \frac{w_{it}}{w_{i\Sigma}} X_{it} \tag{35}$$

and $w_{i\Sigma} = \sum_{t=1}^{n} w_{it}$. Parameters s^2 and a are defined as in equations (29) and (30), respectively.

In practice, the unknown structure parameters m, s^2, and a are estimated from the available data; see [17, 18] or [3] for discussions on this topic. (Function `cm` of the `R` package `actuar` [19] implements all credibility calculations.)

Example 3 The Bühlmann model is appropriate to estimate the pure premiums (or probability of accident) for year 11 in Example 1. Unbiased estimators of the structure parameters are

$$\hat{m} = \bar{S} = 0.23 \tag{36}$$

$$\hat{s}^2 = \frac{1}{I(n-1)} \sum_{i=1}^{I} \sum_{t=1}^{n} (S_{it} - \bar{S}_i)^2 = 0.1367 \tag{37}$$

$$\hat{a} = \frac{1}{I-1} \sum_{i=1}^{I} (\bar{S}_i - \bar{S})^2 - \frac{\hat{s}^2}{n} = 0.0464 \tag{38}$$

Replacing these values in equation (27) yields the estimated credibility premiums of Table 4. One will note that the premium for policyholder 1 obtained with this non parametric approach is different from the pure Bayes premium of Example 2.

Other Credibility Models

Random Coefficients Regression Model

Consider a data set in which a structural linear upward trend is present. Since the Bühlmann and Bühlmann–Straub models assume that the risk premium remains constant in time, these models will

Table 4 Bühlmann credibility premiums for the data of Table 1

Policy holder	Individual premium ($\bar{S}_i$)	Credibility factor ($\hat{z}$)	Credibility premium ($\hat{\pi}_{i,11}$)
1	0.6	0.772	0.516
2	0.3	0.772	0.284
3	0.2	0.772	0.207
4	0.2	0.772	0.207
5	0.2	0.772	0.207
6	0.1	0.772	0.130
7	0.0	0.772	0.052
8	0.0	0.772	0.052
9	0.7	0.772	0.593
10	0.0	0.772	0.052

systematically underestimate future claims. The random coefficients regression credibility model introduced by Hachemeister [20] remedies this situation by relaxing the constant mean assumption of the Bühlmann–Straub model in favor of

$$E[X_{it}|\Theta_i = \theta_i] = \beta_0(\theta_i) + \beta_1(\theta_i)t \qquad (39)$$

Therefore, one has to estimate the intercept and slope of a regression line for each policyholder, but these depend on the unknown risk levels of the policyholders. The optimal estimation of the slope for one policyholder is a credibility weighted average between the estimation using this policyholder's experience only and the estimation using all the portfolio data. The same idea holds for the intercept.

The formulas in this model make heavy use of matrix notation and quickly become cumbersome. We refer the interested reader to [18] for an up-to-date discussion on this topic.

Of course, one can use something other than a linear trend model. In general, one would have

$$E[\mathbf{X}_i|\Theta_i = \theta_i] = \mathbf{Y}\boldsymbol{\beta}(\theta_i) \qquad (40)$$

where $\mathbf{X}_i = (X_{i1}, \ldots, X_{in})'$, $\mathbf{Y}$ is an $n \times (p+1)$ design matrix and $\boldsymbol{\beta}(\theta_i) = (\beta_0(\theta_i), \ldots, \beta_p(\theta_i))'$ is a vector of random coefficients.

Hierarchical Credibility

It is common that the group of policyholders that we have been referring to as *the portfolio* is actually just a fraction of a much larger classification scheme. Consider, for example, a hierarchical (or treelike) structure in which policyholders are classified first in "sectors", and then in "classes". Applying a credibility model within one class will not make use of collateral data from other classes of the same sector, even though this data can be useful to assess the true risk level of the class and, ultimately, of the policyholders (*see* **Dependent Insurance Risks**).

Exploiting collateral data was the main rationale brought forth by Jewell [21] to introduce the hierarchical credibility model. The credibility premium of a policyholder remains of the same form as equation (27) or equation (33), but with the complement of credibility $(1 - z_i)$ given to the *credibility premium* of the class. The latter is, in turn, a weighted average between the experience of the class and the credibility premium of the sector, and so on, until the top level is reached.

Complete formulas for the hierarchical credibility model can be found in [3, 18, 22, 23]. Hierarchical credibility is particularly well suited for a geometrical interpretation; see [24]. Moreover, [3] has a discussion on the uses and misuses of hierarchical models.

Crossed Classification Credibility

Consider an automobile insurance portfolio in which policyholders are classified according to gender and age. Intuitively, young women share risk characteristics with young men, who themselves share risk characteristics with older men. A hierarchical classification structure is inappropriate for such a portfolio since the risk factors are not nested, but rather crossed.

Dannenburg [25] adapted the crossed classification random-effects models, well known in statistical modeling, to the credibility theory context. This results in a very general model encompassing the Bühlmann and Bühlmann–Straub models and, to some extent, the hierarchical model; see [3, 26] for details and formulas.

Concluding Remarks

Credibility theory is related – without being identical – to various forecasting techniques in statistics. There has been a trend over the last decade or so to increasingly draw links between credibility and techniques such as the Kalman filter (*see* **Non-life Loss Reserving**) [27], longitudinal models [28], variance components models [26], or generalized linear models [29]. See also [30] for further comments and references.

A lot of research in credibility theory has been and will likely continue to be devoted to estimation of the structure parameters. Robust estimation or generalized estimating equations come to mind as procedures making their way into credibility. Moreover, the tidal wave of dependence modeling that has swept risk theory around the start of the new millennium has just started to touch credibility theory [31]. More is expected in the near future.

References

[1] Bühlmann, H. (1969). Experience rating and credibility, *ASTIN Bulletin* **5**, 157–165.

[2] Norberg, R. (1979). The credibility approach to ratemaking, *Scandinavian Actuarial Journal* **1979**, 181–221.

[3] Goulet, V. (1998). Principles and application of credibility theory, *Journal of Actuarial Practice* **6**, 5–62.

[4] Neuhaus, W. (2004). Experience rating, in *Encyclopedia of Actuarial Science*, ISBN 0-4708467-6-3, J.L. Teugels & B. Sundt, eds, John Wiley & Sons, Vol. 2, pp. 639–646.

[5] Bühlmann, H. (1967). Experience rating and credibility, *ASTIN Bulletin* **4**, 199–207.

[6] Mowbray, A.H. (1914). How extensive a payroll exposure is necessary to give a dependable pure premium? *Proceedings of the Casualty Actuarial Society* **1**, 25–30.

[7] Longley-Cook, L.H. (1962). An introduction to credibility theory, *Proceedings of the Casualty Actuarial Society* **49**, 194–221.

[8] Whitney, A.W. (1918). The theory of experience rating, *Proceedings of the Casualty Actuarial Society* **4**, 275–293.

[9] Fisher, A. (1919). The theory of experience rating: discussion, *Proceedings of the Casualty Actuarial Society* **5**, 139–145.

[10] Bailey, A.L. (1945). A generalized theory of credibility, *Proceedings of the Casualty Actuarial Society* **32**, 13–20.

[11] Bailey, A.L. (1950). Credibility procedures, Laplace's generalization of Bayes' rule and the combination of collateral knowledge with observed data, *Proceedings of the Casualty Actuarial Society* **37**, 7–23.

[12] Russell, B. (1948). *Human Knowledge: Its Scope and Limits*, George Allen & Unwin, London.

[13] Hogg, R.V., Craig, A.T. & McKean, J.W. (2005). *Introduction to Mathematical Statistics*, ISBN 0-1300850-7-3, 6th Edition, Prentice Hall, Upper Saddle River.

[14] Mayerson, A.L. (1964). A Bayesian view of credibility, *Proceedings of the Casualty Actuarial Society* **51**, 85–104.

[15] Jewell, W.S. (1974). Credible means are exact Bayesian for exponential families, *ASTIN Bulletin* **8**, 77–90.

[16] Bühlmann, H. & Straub, E. (1970). Glaubgwürdigkeit für Schadensätze, *Bulletin of the Swiss Association of Actuaries* **70**, 111–133.

[17] Dubey, A. & Gisler, A. (1981). On parameter estimation in credibility, *Bulletin of the Swiss Association of Actuaries* **81**, 187–211.

[18] Bühlmann, H. & Gisler, A. (2005). *A Course in Credibility Theory and Its Applications*, ISBN 3-5402575-3-5, Springer.

[19] Goulet, V. (2005). *Actuar: an R Package for Actuarial Science*, http://www.actuar-project.org.

[20] Hachemeister, C.A. (1975). Credibility for regression models with application to trend, in *Credibility, Theory and Applications*, *Proceedings of the Berkeley Actuarial Research Conference on Credibility*, Academic Press, New York, pp. 129–163.

[21] Jewell, W.S. (1975). The use of collateral data in credibility theory: a hierarchical model, *Giornale dell'Istituto Italiano degli Attuari* **38**, 1–16.

[22] Goovaerts, M.J., Kaas, R., van Heerwaarden, A.E. & Bauwelinckx, T. (1990). *Effective Actuarial Methods*, ISBN 0-4448839-9-1, North-Holland, Amsterdam.

[23] Goovaerts, M.J. & Hoogstad, W.J. (1987). *Credibility Theory*, *Surveys of Actuarial Studies*, Number 4, Nationale-Nederlanden, Netherlands.

[24] Bühlmann, H. & Jewell, W.S. (1987). Hierarchical credibility revisited, *Bulletin of the Swiss Association of Actuaries* **87**, 35–54.

[25] Dannenburg, D. (1995). Crossed classification credibility models, in *Transactions of the 25th International Congress of Actuaries*, International Actuarial Association Vol. 4, pp. 1–35.

[26] Dannenburg, D.R., Kaas, R. & Goovaerts, M.J. (1996). *Practical Actuarial Credibility Models*, ISBN 9-0802117-3-7, Ceuterick, Leuven.

[27] de Jong, P. & Zehnwirth, B. (1983). Credibility theory and the Kalman filter, *Insurance: Mathematics and Economics* **2**, 281–286.

[28] Frees, E.W., Young, V. & Luo, Y. (1999). A longitudinal data analysis interpretation of credibility models, *Insurance: Mathematics and Economics* **24**, 229–247.

[29] Kaas, R., Goovaerts, M., Dhaene, J. & Denuit, M. (2001). *Modern Actuarial Risk Theory*, ISBN 0-7923763-6-6, Kluwer Academic Publishers, Dordrecht.

[30] Norberg, R. (2004). Credibility theory, in *Encyclopedia of Actuarial Science*, ISBN 0-4708467-6-3, J.L. Teugels & B. Sundt, eds, John Wiley & Sons, Vol. 1, pp. 398–406.

[31] Denuit, M., Dhaene, J., Goovaerts, M. & Kaas, R. (2005). *Actuarial Theory for Dependent Risks*, ISBN 0-4700149-2-X, John Wiley & Sons, Chichester.

VINCENT GOULET