A new approach to the credibility formula

Amir T. Payandeh Najafabadi

Department of Mathematical Sciences, Shahid Beheshti University, G.C. Evin, 1983963113, Tehran, Iran

2010

Online at https://mpra.ub.uni-muenchen.de/21587/
MPRA Paper No. 21587, posted 13. April 2010 02:29 UTC
A new approach to the credibility formula

AMIR T. PAYANDEH NAJAFABADI

a Department of Mathematical Sciences, Shahid Beheshti University,
G.C. Evin, 1983963113, Tehran, IRAN.
Email: amirtpayandeh@sbu.ac.ir

ABSTRACT

The usual credibility formula holds whenever, (i) claim size distribution is a member of the exponential family of distributions, (ii) prior distribution conjugates with claim size distribution, and (iii) square error loss has been considered. As long as, one of these conditions is violated, the usual credibility formula is no longer hold. This paper using the mean square error minimization technique develops a simple and practical approach to the credibility theory. Namely, we approximate the Bayes estimator with respect to a general loss function and general prior distribution by a convex combination of the observation mean and mean of prior, say, approximate credibility formula. Adjustment of the approximate credibility for several situations and its form for several important losses are given.

JEL Classification: C11, C16

Keywords: IM31, loss function, balanced loss function, mean square error technique.

1. Introduction

Credibility theory is the art of combining different collections of data to obtain an accurate overall estimate. It provides actuaries with techniques to determine insurance premiums for contracts that belong to a (more or less) heterogeneous portfolio, where is limited or irregular claim experience for

1\footnote{August 9, 2009}
each contract but ample claim experience for the portfolio. Credibility theory can be seen as one of quantitative tools that allows the insurers to perform experience rating, that is, to adjust future premiums based on past experiences. In many cases, a compromise estimator is derived from a convex combination of prior mean and the mean of the current observations. Credibility theory began with papers by Mowbray (1914) and Whitney (1918). These papers purposed to derive the new premium with a convex combination of prior mean, say, \( \mu \) and the mean of the current observations, say, \( \bar{X} \) by \( P = z \bar{X} + (1-z)\mu \), where \( z \) represents the credibility factor, ranging from 0 to 1. Bailey (1950) showed that the formula \( P = z \bar{X} + (1-z)\mu \) may be derived from Bayes’ theorem, either by using a Bernoulli-Beta model on unknown parameter \( p \), or by the using a Poisson-Gamma model on unknown parameter \( \lambda \). Bailey’s work led to the application of Bayesian methodology to credibility theory. Excellent introduction to credibility theory can be found, e.g., in Goovaerts & Hoogstad (1996), Herzog (1994), Dannenburg, Kass, & Goovaerts (1996), Klugman, Panjer, & Willmot (2004, Chapter 16) and Bühlmann & Gisler (2005). See also Norberg (2004) for an overview with useful references and links to Bayesian statistics and linear estimation. There are many applications of credibility techniques to various branches of insurance. Rejesus, Coble, Knight, & Jin (2006) provided a nonstandard application of credibility techniques. Namely, they examined the feasibility of implementing and experience-based premium rate discount in crop insurance.

However, the credibility restricted by family of distributions, conjugate prior, and square error loss functions. Neither the claim distributions which are not members of the exponential family of distributions nor the non-conjugate prior, the predicted mean (Bayes estimator with respect to square error loss) is no longer linear with respect to the data (see Diaconis & Ylvisaker 1997) and the credibility formula is no longer true. Whenever the policyholder is undercharged (and insurance company loses its money) or the insured is overcharged (and the insurer is at risk of losing the policy), the square loss assigns similar penalty to over- and undercharge. In order to assign more (or less) penalty to overcharged, one has to consider a loss function rather than square error loss to reflect such concerns. Loss functions rather than square error and Entropy losses, usually, leads to a Bayes estimator
which cannot be a linear combination of observation mean and mean of prior distribution. Therefore, in such cases, the credibility formula is no longer hold. Bühlmann (1967) overcame the prior limitation and proved that in class of linear estimators with form \( \delta_{Lin}(X_1, \cdots, X_n) = c_0 + \sum_{j=1}^{n} c_j X_j \), an estimator \( P = z\bar{X} + (1-z)\mu \), is also a distribution free credibility formula, which minimizes \( E\{\mu(\theta) - \delta_{Lin}(X_1, \cdots, X_n)\}^2 \), whenever \( \mu(\theta) \) is the mean of an individual risk (or \( \mu(\theta) = E(X|\theta) \)), characterized by risk parameter \( \theta \), and \( \bar{X} = (X_1 + X_2 + \cdots + X_n)/n \). Bühlmann & Straub (1970) then formalized the least squares derivation of \( z = n/(n + k) \), where \( n \) is the number of trials or exposure units and \( k = v/a \), in which \( v = E(Var(X|\theta)) \) and \( a = Var(E(X|\theta)) \). This new methodology is called empirical Bayes credibility, although the Bayesian content of this approach has been greatly minimized. Landsman (2002) used the second order statistical technique and established a new approach to the credibility theory. Bühlmann (1967), Bühlmann & Straub (1970), and Landsman (2002) overcame restriction of the credibility theory by the exponential family of distinctions and the conjugate priors. But, restriction by the square error loss function, still, is reminded. An extension to the credibility formula from the loss viewpoint given by Gomez (2006, 2007). He considered the balanced square loss function, \( L_{p,\omega,\delta_0}(\theta, \delta) = \omega h(\theta)(\delta - \delta_0)^2 + (1-\omega)h(\theta)(\delta - \theta)^2 \), where \( \delta_0 \) is chosen a prior "target" estimator of \( \theta \), obtained for instance from the criterion of maximum likelihood estimator, least-squares, or unbiased among others. He established the credibility theory for this loss function.

This paper develops a simple and applicable credibility formula which is obtained by approximating the Bayes estimator by a convex combination of observation mean and mean of prior, say, approximate credibility formula. Adjustment of the approximate credibility for several situations and its form for several important losses are given.

This paper develops as the following. Section 2 collects some useful elements for other sections. Section 3 provides a new approach to credibility formula by establishing an approximate credibility premium, say approximate credibility formula. Adjustment of the approximate credibility for several important situations and for several important losses are discussed.
2. Preliminaries

It is useful to recall that, family of densities function \( \{ p_{\theta}(\cdot) : \theta \in \Theta \} \) is said to have monotone likelihood ratio (mlr) in \( T(\cdot) \), such that for all \( \theta_1 > \theta_2 \), the densities \( p_{\theta_i}(\cdot) \), for \( i = 1, 2 \), are distinct, and ratio \( p_{\theta_1}(x)/p_{\theta_2}(x) \) is a nondecreasing function of \( T(x) \). The following from Lehmann & Romano (2005) recalls an important property of a class of density functions which have the mlr property.

**Lemma 1.** *(Kline-Rubin’s lemma)* Suppose \( \{ p_{\theta}(\cdot) : \theta \in \Theta \} \) is a family of density functions with the mlr in \( x \). Moreover, suppose that \( \psi(x) \) is a nondecreasing function in \( x \). Then \( E(\psi(X)|\theta) \) is nondecreasing function in \( \theta \).

Now, we recall definition of symmetric log-concave density functions.

**Definition 1.** Random variable \( X \), given location parameter \( \theta \), has symmetric log-concave density function \( f_0 \) if and only if \( f_0(x, \theta) \propto \exp(-k(x - \theta)) \), where \( k \) is a function in class of functions \( H^* := \{ k : k \text{ symmetric about zero, increasing, convex, and } k' \text{ concave} \} \).

When the new premium amount is fixed by insurer, two kinds of errors can be arisen: either the policyholder is undercharged and insurance company loses its money or the insured is overcharged and the insurer is at risk of losing the policy. In order to penalize large mistakes to a greater extent, it is usual to consider a nonnegative convex function as a loss function. The loss function is generally taken to be square error loss, which gives same penalty to undercharge and overcharge. But in many cases, we may be interested to loss functions which assigned more (less) penalty to overcharges. In decision theory, Entropy loss function (given by \( L_{Ent}(\theta, \delta) = \theta/\delta - \ln(\theta/\delta) - 1 \)) and Linex loss function (given by \( L_{Linex}(\theta, \delta) = \exp\{a(\delta - \theta)\} - a(\delta - \theta) - 1 \) with \( a > 0 \)) are two popular losses which consider in situation that overestimation is more considerable than underestimation. Meanwhile, in the reverse situation (underestimation is more considerable than overestimation) Stein loss function (given by \( L_{Stein}(\theta, \delta) = \delta/\theta - \ln(\delta/\theta) - 1 \)) and Linex loss function (given by \( L_{Linex}(\theta, \delta) = \exp\{a(\delta - \theta)\} - a(\delta - \theta) - 1 \) with \( a < 0 \)) are more applicable losses.
In existence of an estimator (say target estimator) with some useful properties, such as admissibility, maximum likelihood, minimaxity, etc., we may interest to class of loss functions which involve the target estimator to find a good (in some sense) estimator. Zellner (1994) introduced a class of loss functions, named balanced losses, which gives weight of \( \omega \) to penalty of distance from target estimator \( \delta_0 \) and weight of \( 1 - \omega \) to distance from true parameter \( \theta \). Dey, Ghosh, & Strawderman (1999) generalized Zellner’s balanced loss function to a class of balanced loss function with form

\[
L_{\rho, \omega, \delta_0}(\theta, \delta) = \omega h(\theta) \rho(\delta_0, \delta) + (1 - \omega) h(\theta) \rho(\theta, \delta),
\]

where \( \delta_0 \) is chosen a prior "target" estimator of \( \theta \). Several issues, such as Bayesianity, admissibility, dominance, and minimaxity studied by Jafari Jozani, Marchand, & Parsian (2006).

The following lemma explores an important property of Bayes estimator.

**Lemma 2.** Suppose random variable \( X \) given location parameter \( \theta \) has been distributed according to a symmetric log-concave density function, given by Definition (1). Then, Bayes estimators with respect to prior \( \pi(\theta) \) and square error, Entropy, Stein, and Linex loss functions are nondecreasing in \( x \).

**Proof:** Despite theoretical differences between square error and Entropy losses, Bayes estimator under both loss functions given by posterior mean \( E_\pi(\theta|x) \); while Bayes estimator under Stein and Linex losses are given by \( 1/E_\pi(1/\theta|x) \) and \(-\ln(E_\pi(\exp\{-a\theta\}|x))/a \), respectively. These observations along Lemma (1) and the fact that posterior distribution \( \theta \) given \( x \) has the mlr property in \( \theta \), whenever \( x \) viewed as a parameter, complete the desire proof. \( \Box \)

### 3. Main results

It is well known that Bayes estimator reflects properties of loss function and prior distribution (see Payandeh & Marchand 2009). Therefore, it makes sense to consider a Bayes estimator, under an appropriate loss and prior distribution, as a suitable and acceptable estimator which reflects our
concerns about an unknown parameter and biasness of an estimator. The following lemma considers a Bayes estimator as an appropriate estimator for parameter \( \theta \). Then using the mean square error technique develops a new approach to approximate the Bayes estimator by the credibility formula.

**Lemma 3.** Suppose \( X_1, X_2, \cdots, X_n \) given risk parameter \( \theta \) are identical and independent distributed with \( \mu(\theta) = E(X|\theta) \), for \( i = 1, 2, \cdots, n \). Moreover, suppose that risk parameter \( \theta \) has prior distribution \( \pi \) with mean \( \mu \), i.e., \( \mu = E_{\pi}(\theta) \), and \( \delta_{\pi} \) is a Bayes estimator with respect to loss function \( \rho \) and prior distribution \( \pi \). Then, in the class of credibility premiums

\[
\delta = \{ \delta_{\alpha} : \text{ where } \delta_{\alpha}(\bar{x}) = \alpha \bar{x} + (1 - \alpha)\mu \text{ and } \alpha \in [0, 1] \}
\]

an estimator \( \delta_{opt} \), with

\[
\alpha_{opt} = \frac{E((\bar{X} - \mu)(\delta_{\pi}(X) - \mu))}{E((\bar{X} - \mu)^2)};
\]

minimizes the mean squared error between \( \delta_{\pi} \) and \( \delta_{\alpha} \), i.e., \( \delta_{opt} = \arg \min E(\delta_{\pi}(X) - \delta_{\alpha}(X))^2 \), where \( \bar{X} = (X_1, X_2, \cdots, X_n)^T \) and two-folded expectation \( E(\cdot) \) stands for \( E_{\pi}(E(\cdot|\theta)) \).

**Proof:** Mean square distance between two estimators \( \delta_{\pi} \) and \( \delta_{\alpha} \) can be readily observed as

\[
\text{MSE}(\alpha) = E(\delta_{\pi}(X) - \delta_{\alpha}(X))^2 = E(\delta_{\pi}(X) - \alpha \bar{X} - (1 - \alpha)\mu)^2.
\]

Taking derivative with respect to \( \alpha \) along the fact that second derivative of MSE(\( \alpha \)) with respect to \( \alpha \), \( \text{MSE}''(\alpha) = 2E(\bar{X} - \mu)^2 \), is nonnegative lead to desire result.

Two-folded expectations in nominator and denominator of \( \alpha_{opt} \) given the above can be simplified as

\[
E((\bar{X} - \mu)(\delta_{\pi}(X) - \mu)) = E_{\pi}(Cov(\bar{X}, \delta_{\pi}(X)|\theta)) + Cov_{\pi}(E(\bar{X}|\theta), E(\delta_{\pi}(X)|\theta))
\]

\[
+ (\mu_0 - \mu)(\mu_{\delta_{\pi}} - \mu);
\]

\[
E(\bar{X} - \mu)^2 = E_{\pi}(Var(\bar{X}|\theta)) + Var_{\pi}(E(\bar{X}|\theta)) + (\mu_0 - \mu)^2,
\]

\[6\]
where $\mu = E_\pi(\theta)$, $\mu^\delta = E_\pi(E(\delta(X)|\theta))$, and $\mu_0 = E_\pi(E(X_i|\theta))$, for $i = 1, 2, \cdots, n$. The following remark summarizes the above observation along a double applications of the Wald’s identity for conditional covariance

**Remark 1.** Under conditions given by Lemma 3, the $\alpha_{opt}$ may be represented as

$$
\alpha_{opt} = \frac{Cov(\bar{X}, \delta_{\pi}(X)) + (\mu_0 - \mu)(\mu^\delta - \mu)}{Var(\bar{X}) + (\mu_0 - \mu)^2}.
$$

It worth to mention that the optimal $\alpha$ given by Lemma 3 is applicable, whenever $0 \leq \alpha_{opt} \leq 1$. In the case that $\alpha_{opt}$ exceed interval $[0, 1]$, it can be modified by projecting into interval $[0, 1]$.

An estimator $\delta_{opt}$ in Lemma 3 can be criticized, because it gives the claim amounts for all previous years the same weight; intuitively one should believe that new claims should have more weight than old claims. However, as the claim amounts given the risk parameter of different years were assumed to be exchangeable, it was only reasonable that the claim amounts given the risk parameter should be identically distributed but the risk parameter varies for each year. Atansiu (2008) considered a model in which the risk parameters for $n$ years have joint prior density $\pi(\theta_1, \cdots, \theta_n)$. To reflect the fact that the correlation between claim amount for different years has to decrease as the time distance between the years increase, he considered model with assumption that

$$
Cov(E(X_i|\theta_i), E(X_j|\theta_j)) = \rho^{|i-j|}\lambda,
$$

where $0 < \rho < 1$, and $\lambda > 0$. Then, he suggest to use $\alpha_0 + \sum_{j=1}^{t} \alpha_j X_j$, where $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_t < 1$, as the credibility premium. Same as Atansiu (2008) in such situations, we suggest to replace the arithmetic mean $\bar{X}$ by the weighted mean $\bar{X}_{\omega} = \sum_{j=1}^{t} \omega_j X_j$, where $0 < \omega_1 < \omega_2 < \cdots < \omega_t < 1$ and $\sum_{j=1}^{t} \omega_j = 1$, in Lemma 3.

Often, in particular in reinsurance, one wants to allow for varying risk volumes, and for that purpose we will introduce the credibility model incorporating risk volumes. We consider a ceded insurance portfolio. Suppose that the claim amounts $Y_{i_1}^j, Y_{i_2}^j, \cdots, Y_{i_m}^j$ of the risks in year $j$, where $m_j$ some measure of the risk volume in year $j$. By the loss ratio of year $j$, we shall mean $X_j = S_j/m_j$, where
Same as Atansiu (2008) in such cases, we suggest to replace $\bar{X}$ in Lemma 3 by $\sum_{j=1}^{n} X_j / \sum_{j=1}^{n} m_j$.

Results given by Lemma 3 are valid for all kinds of risk parameter. But it is difficult to establish $\alpha_{opt}$ given by Lemma 3 lays in a interval $[0, 1]$. Therefore, from here to end of this paper, we study especial case of the location risk parameter. Now, suppose risk parameter, $\theta$, is a location parameter and claim size random variable, $X$, given risk parameter, $\theta$, has been distributed according to a symmetric log-concave density function, given by Definition 1. In situations where the exact credibility premium is not hold, the next theorem, using Lemma 3 provides an approximate credibility premium.

**Theorem 1.** Suppose claim size random variables $X_1, X_2, \cdots, X_n$, given location risk parameter $\theta$, randomly sampled from a symmetric log-concave density function, Definition 1. Moreover, suppose that risk parameter $\theta$ has prior distribution $\pi$. Then, credibility factor of the approximate credibility premium, given in Lemma 3: (i) simplified to $\alpha_{opt} = [E_{\pi}(\text{Cov}(\bar{X}, \delta_{\pi}(\tilde{X})|\theta)) + \text{Cov}_{\pi}(\theta, E(\delta_{\pi}(\tilde{X})|\theta))]/[\text{Var}(\bar{X}|\theta)/n + \text{Var}_{\pi}(\theta)]$, (ii) $0 \leq \alpha_{opt} \leq 1$, whenever $\delta_{\pi}$ is Bayes estimator, with respect to one of square error, Entropy, Stein, or Linex loss functions.

**Proof.** Proof (i) obtains from the fact that, for all $i = 0, 1, \cdots, n, \mu_0 := E_{\pi}(E(X_i|\theta)) = E_{\pi}(\theta) =: \mu$. For (ii) observe that Bayes estimator, under square error, Entropy, Stein, or Linex losses is an increasing function in $x$, see Lemma 2, this observation along the fact that covariance between two nondecreasing functions is nonnegative establish nonnegativity of $\alpha_{opt}$. To establish $\alpha_{opt} \leq 1$, from Remark 1 observe that $\alpha_{opt} = \text{Cov}(\bar{X}, \delta_{\pi}(\tilde{X}))/\text{Var}(\bar{X})$. Now, recall that $\text{Cov}(\bar{X}, \delta_{\pi}(\tilde{X}))$ maximizes whenever $\delta_{\pi}(\bar{x}) = a + b\bar{x}$. Recent observation valid, whenever the exact credibility premium holds. Therefore, $b \leq 1$, and consequently $\alpha_{opt} \leq 1$. □

The above theorem provides an approximate credibility premium in situations that the exact credibility premium does not hold. The natural question that arises is that: in the existence of the exact credibility formula, how the approximate credibility behaves? The following explores this case.

**Lemma 4.** In the existence of exact credibility premium the approximate credibility premium, given
by Theorem 1, coincides with exact one.

**Proof:** If the exact credibility premium holds the Bayes estimator \( \delta_n \) can be written as \( \delta_n(x) = z\bar{x} + (1 - z)\mu \). From this fact, one can observe that

\[
\alpha_{opt} = \frac{E((\bar{x} - \mu)(\delta_\pi(X) - \mu))}{E((\bar{x} - \mu)^2)} = \frac{E((\bar{x} - \mu)(z\bar{x} + (1 - z)\mu - \mu))}{E((\bar{x} - \mu)^2)} = z. \quad \square
\]

The following provides an example to realize application of the approximate credibility premium in such situations, where the exact credibility premium does not applicable.

**Example 1.** Suppose \( X|\theta \) has been distributed according to Normal distribution with mean \( \theta \) variance \( \sigma^2 = 1 \), and unknown parameter \( \theta \) distributed according to non-conjugate prior Gamma(2,2). The approximate credibility premiums, respectively, for square error, Entropy, Stein, and Linex (with \( a = 0.2 \) and \( a = -0.2 \)) losses are

- \( \delta_{\text{Square-error}}^{\text{opt}}(x) = 0.2395\bar{x} + 0.7605 \)
- \( \delta_{\text{Entropy}}^{\text{opt}}(x) = 0.2395\bar{x} + 0.7605 \)
- \( \delta_{\text{Stein}}^{\text{opt}}(x) = 0.36475\bar{x} + 0.6353 \)
- \( \delta_{\text{Linex},a=0.2}^{\text{opt}}(x) = 0.1889\bar{x} + 0.8111 \)
- \( \delta_{\text{Linex},a=-0.2}^{\text{opt}}(x) = 0.1791\bar{x} + 0.8209 \)

Figures 1 compares risk of these approximate credibility premiums with their corresponding Bayes estimators for such loss functions. As all figures show: (i) for small value \( \theta \) the approximate credibility premiums are closed to Bayes estimators, (ii) in some intervals the approximate credibility estimator performance better that the Bayes one.

Gomez (2006, 2007) established that, we may have the exact credibility premium for weighted balanced square error loss functions. Gomez’s result can readily extent to weighted balanced entropy loss function. The next lemma establishes the approximate credibility formula for weighted balanced square error and entropy losses.

**Theorem 2.** Suppose claim size random variables \( X_1, X_2, \cdots, X_n \), given location risk parameter \( \theta \), Definition 1. Moreover, suppose that risk parameter \( \theta \) has prior distribution \( \pi \), and \( \delta_0(x) \) is a
target estimator which nondecreasing in \( x \), which \( \text{Var}(\bar{X}) \leq \text{Var}(\delta_0(X)) \). Then, credibility factor of the approximate credibility premium, given in Lemma 3: (i) simplified to

\[
\alpha_{\text{opt}}^\omega = \omega \frac{E_\pi(Cov(\bar{X}, \delta_\pi(X)|\theta)) + Cov_\pi(\theta, E(\delta_0(X)|\theta))}{\text{Var}(X|\theta)/n + Var_\pi(\theta)} + (1 - \omega) \frac{E_\pi(Cov(\bar{X}, \delta_\pi(X)|\theta)) + Cov_\pi(\theta, E(\delta_\pi(X)|\theta))}{\text{Var}(X|\theta)/n + Var_\pi(\theta)},
\]

(ii) \( 0 \leq \alpha_{\text{opt}}^\omega \leq 1 \), where \( \delta_\pi \) is the Bayes estimator with respect to the balanced square error or entropy losses.

**Proof:** Part (i) obtains after a straightforward calculation. Nonnegativity of \( \alpha_{\text{opt}}^\omega \), in part (ii), follows from nondecreasing in \( x \) of \( \delta_0(\bar{x}) \) and \( \delta_\pi(\bar{x}) \) along the fact that covariance between non-decreasing functions is nonnegative. To establish \( \alpha_{\text{opt}}^\omega \leq 1 \) from assumptions on \( \delta_0 \) observe that \( \text{Cov}(\bar{X}, \delta_0(X)) \leq \text{Var}(\bar{X}) \) and with a similar argument with Theorem 1 observe \( \text{Cov}(\bar{X}, \delta_\pi(X)) \leq \text{Var}(\bar{X}) \). Now, an application of Remark 1 completes desire proof.

4. Conclusion

This paper provides a technique to chose credibility factor \( \alpha_{\text{opt}} \) such that Bayes estimator \( \delta_\pi \), under an appropriate loss and prior distribution, can be approximated by \( \alpha_{\text{opt}} \bar{x} + (1 - \alpha_{\text{opt}})\mu \). Definitely, to use such approximation, one has to establish \( \alpha_{\text{opt}} \) lays in an interval \([0, 1]\). For a family of symmetric logconcave distributions with location parameter, under Square error, Entropy, Stein, and Balanced loss functions, the above requirement on \( \alpha_{\text{opt}} \) has been established. It is worth to mention that the idea that develops by Lemma 3 may be employed: (i) For general risk parameter; (ii) The credibility premium in general setting, for instance: credibility for the chain ladder reserving method (Gisler & Wüthrich 2008) and credibility premiums for the zero-inflated Poisson model (Bouchera & Denuit 2008), among others.
References


