Annals of Actuarial Science (2026), **20**, pp. 7–21
doi:10.1017/S1748499525000065

ORIGINAL RESEARCH PAPER

Mixture credibility formulas

Mojtaba Abed and Amir T. Payandeh Najafabadi

Department of Actuarial Science, Faculty of Mathematical Sciences, Shahid Beheshti University, Evin Tehran, Iran
Corresponding author: Amir T. Payandeh Najafabadi; Email: amirtpayandeh@sbu.ac.ir

(Received 31 May 2024; revised 04 April 2025; accepted 12 April 2025; first published online 19 June 2025)

Abstract

The classical credibility premium provides a simple and efficient method for predicting future damages and losses. However, when dealing with a nonhomogeneous population, this widely used technique has been challenged by the Regression Tree Credibility (RTC) model and the Logistic Regression Credibility (LRC) model. This article introduces the Mixture Credibility Formula (MCF), which represents a convex combination of the classical credibility premiums of several homogeneous subpopulations derived from the original population. We also compare the performance of the MCF method with the RTC and LRC methods. Our analysis demonstrates that the MCF method consistently outperforms these approaches in terms of the quadratic loss function, highlighting its effectiveness in refining insurance premium calculations and enhancing risk assessment strategies.

Keywords: Bayesian credibility mean; classification; credibility formula; dominating estimators; squared error loss function

1. Introduction

Consider a sequence of independent and identically distributed (i.i.d.) random variables $X_1, \ldots, X_n$, each sampled from a population characterized by a risk parameter θ. The Bühlmann credibility, denoted as $P^{B.C.}$, for estimating $E(X_{n+1}|X_1, X_2, \ldots, X_n)$ is expressed as a convex combination of the collective premium, $\bar{X}$ and the individual premium μ_0, that is, $P^{B.C.} = \zeta \bar{X} + (1-\zeta)\mu_0$, where the credibility factor ζ is defined as $\zeta = \frac{n}{n+\sigma^2/\tau^2}$, with $\sigma^2 = E\left[Var(X_1|\theta)\right]$ and $\tau^2 = Var\left[E(X_1|\theta)\right]$.

When significant heterogeneity is present within the population, the ratio σ^2/τ^2 tends toward infinity, causing the credibility factor ζ to approach zero. The fundamental task in risk rating is the determination of the so-called pure risk premium $E[X_i]$. Such that the pure, or collective, premium is the mean of the hypothetical means. This is the premium we would use if we knew nothing about the individual. It does not depend on the individual's risk parameter, θ, nor does it utilize $\boldsymbol{X} = \boldsymbol{x}$, the data collected from the individual. Because θ is unknown, the best we can do is leverage the available data, which suggest the use of the Bayesian premium (the mean of the predictive distribution) $E(X_{n+1}|X_1, X_2, \ldots, X_n)$; see Bühlmann and Gisler (2005) and Kass *et al.* (2008) for more details. Various authors have addressed this issue and proposed potential solutions. For example, Diao and Weng (2019) introduced the Regression Tree Credibility (RTC) model, while Jahanbani *et al.* (2024) proposed the Logistic Regression Credibility (LRC) model.

The roots of credibility theory trace back to the seminal works of Mowbray (1914) and Whitney (1918), who proposed a convex combination of collective and individual premiums as the optimal insurance contract premium. Bailey (1950) formalized this concept, known as the exact credibility

© The Author(s), 2025. Published by Cambridge University Press on behalf of Institute and Faculty of Actuaries. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S1748499525000065 Published online by Cambridge University Press

premium, within the framework of parametric Bayesian statistics. Building upon this foundation, Bühlmann (1967) and Bühlmann and Straub (1970) expanded the notion of the exact credibility premium to a model-based approach, which significantly contributed to the widespread adoption of credibility theory across various actuarial domains.

There are two distinct approaches to credibility theory: European and American; see Norberg (1979) for more details. The European approach, often termed Bayesian credibility theory, prioritizes the integration of prior knowledge into the estimation process. It employs Bayesian statistics to incorporate prior beliefs about an insured individual's risk profile into the prediction of future claims experience. Conversely, the American approach, known for its focus on the credibility premium, relies heavily on past claims data to forecast future losses. Utilizing statistical methods such as the Bühlmann credibility formula, it calculates a credibility factor based on historical data, thereby adjusting the individual's premium accordingly. Although both approaches aim to enhance risk assessment and premium determination in insurance, they differ in their reliance on prior knowledge and historical data, reflecting distinct methodologies within actuarial practice.

For an in-depth exploration of the evolution and methodologies within credibility theory, Bühlmann and Gisler (2005) and Payandeh Najafabadi (2010) provide comprehensive discussions. While classical credibility theory offers a straightforward yet somewhat rigid approach to predictive distribution, for example, Hong and Martin's research (2017, 2018) introduced flexible Dirichlet process mixture models for predicting insurance claims and analyzing loss data. Their work demonstrated the effectiveness and adaptability of Bayesian nonparametric frameworks compared to traditional parametric methods. Their research delved into the theoretical underpinnings and benefits of this approach, comparing it with classical credibility theory. Cheung *et al.* (2022) and Yong *et al.* (2024) offered innovative credibility-based approaches to address challenges in risk estimation and decision-making, integrating concepts from Bayesian hierarchical modeling, prospect theory, and variance premium principle to enhance the accuracy and practicality of actuarial processes. More recently, Gómez-Déniz and Vázquez-Polo (2022) presented a method for deriving exact credibility reference Bayesian premiums based on prior distributions constructed from available data and a generated model, offering a practical solution when prior information is insufficient for premium determination in credibility theory.

Further advancements in Bayesian credibility theory have focused on the selection of prior distributions, as investigated by Hong and Martin (2022), highlighting the importance of precise prior specification in Bayesian credibility modeling. The Bayesian credibility mean under mixture distributions has garnered attention from several researchers, including Lau *et al.* (2006), Cai *et al.* (2015), Hong and Martin (2017, 2018), Zhang *et al.* (2018), Payandeh Najafabdi and Sakizadeh (2019, 2024), Li *et al.* (2021), and others. However, many of these approaches rely on approximations. For example:

(1) Payandeh Najafabadi and Sakizadeh (2019) employed a mixture distribution to approximate the complex posterior distribution, subsequently deriving an approximation for the Bayesian credibility means. Unfortunately, their approximation error increases with the number of past experiences.

(2) Following Lo (1984) and Lau *et al.* (2006) reformulated the predictive distribution of X_{n+1} given past claim experiences $X_1, X_2, \ldots, X_n$ as a finite sum over all possible partitions of the past claim experiences. They then utilized the credibility premium, a convex combination of the collective premium (prior mean) and the sample average of past claim experiences, to derive the Bayesian credibility mean. Notably, within the exponential family of distributions, such a credibility premium coincides with the Bayesian credibility mean, as detailed by Payandeh Najafabadi (2010) and Payandeh Najafabadi *et al.* (2012).

On the contrary, the American credibility (for the sake of convince, just say credibility) theory holds a foundational position within actuarial science, recognized as a cornerstone of insurance

https://doi.org/10.1017/S1748499525000065 Published online by Cambridge University Press

experience rating (Hickman & Heacox, 1999) and tracing its origins back to Whitney (1918). This theory conceptualizes the net premium of individual risk, denoted as $\mu(\Theta)$, as a function of a random element Θ, representing the unobservable characteristics of the individual risk. The (credibility) premium is then calculated as a linear combination of the average rate of individual claims experience and the collective net premium. Credibility theory has evolved into two main streams: limited fluctuation credibility theory and greatest accuracy credibility theory. The former prioritizes stability, aiming to incorporate individual claims experience while maintaining premium stability. Conversely, the latter, widely employed in modern applications, focuses on achieving the minimum mean squared prediction error by utilizing the best linear unbiased estimator to approximate individual net premiums. This approach integrates both individual and collective claims experience to optimize prediction accuracy.

This article proceeds under the assumption that through the use of an appropriate classification method, a random sample of observations $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ can be effectively categorized into k distinct homogeneous subpopulations $\boldsymbol{X}_l = (X_1, X_2, \ldots, X_{n_l})$ for $l = 1, 2, \ldots, k$ (n_l is the number of observations that fall into class l). It then introduces the expression:

$$P^{\mathrm{MCF}} = \sum_{l=1}^{k} \omega_l P_l^{\mathrm{Cred}} = \sum_{l=1}^{k} \omega_l \{\alpha_l \bar{X}_l + (1 - \alpha_l)\mu_l\},$$

where ω_l are premium mixture weights obtained based on a statistical technique, such as the logistic regression, such that $\sum_{l=1}^{k} \omega_l = 1$, α_l credibility weight for class l as $\alpha_l = n_l/(n_l + \sigma_l^2/\tau_l^2)$, μ_l, is collective premium of class l and $\bar{X}_l = \frac{1}{n_l} \sum_{i=1}^{n_l} x_i$ for $l = 1, 2, \ldots, k$.

The rest of this article unfolds as follows: Section 2 elucidates our model assumptions, provides an overview of the Bayesian credibility model, explores the advantages of data space partitioning, and outlines a general credibility model based on partitioning. Section 3 delineates the process of formulating a Mixture Credibility Formula (MCF) and explains the procedure for premium prediction calculations. Section 4 applies our MCF-based prediction model to real-world Medicare data. Finally, Section 5 summarizes the results of the article.

2. Preliminaries and model assumptions

In classical insurance analysis, the treatment of insurance claims typically revolves around modeling them as random variables, denoted as X, with corresponding density functions represented by $f(x|\theta)$. Here, θ signifies a fixed parameter associated with the risk, though often obscured by uncertainty. However, the Bayesian approach to risk analysis introduces a fundamental shift in perspective: θ is no longer regarded as a fixed value but rather as a random variable itself. This paradigmatic change necessitates the adoption of prior distributions for Θ within heterogeneous insurance portfolios.

The Bayesian framework offers considerable advantages to actuaries. First, it liberates them from the constraints of specific models, fostering adaptability in tackling a wide array of problems. Once adept at analyzing one scenario, actuaries find themselves equipped to handle analogous, albeit more intricate, situations with minimal additional complexity. This inherent flexibility is invaluable in navigating the dynamic landscape of insurance and risk management.

Second, the Bayesian approach simplifies the process of obtaining estimates. Notably, it provides a seamless transition from obtaining point estimates to deriving interval estimates. Also, the key point in the Bayesian approach is it allows actuaries to assess the uncertainty of their inference in terms of probabilities. This characteristic is particularly pertinent in contemporary actuarial practice, where stakeholders increasingly demand not only precise estimates but also insights into the uncertainty surrounding them.

https://doi.org/10.1017/S1748499525000065 Published online by Cambridge University Press

By embracing the Bayesian framework, actuaries can navigate the complexities of risk analysis with greater confidence and efficiency. It empowers them to provide robust evidence regarding the quality and reliability of their estimates, thereby enhancing decision-making processes within the insurance and financial industries. As the landscape of risk management continues to evolve, the Bayesian approach stands as a beacon of innovation and adaptability in the pursuit of informed decision-making and risk mitigation strategies.

Consider a scenario where $X_1, X_2, \ldots, X_n$ represent a vector of insurance claims, while $Z_i = (Z_{i1}, Z_{i2}, \ldots, Z_{im})$ denotes the covariate vector associated with individual risk θ_i, for $i = 1, 2, \ldots, n$. Each individual risk is characterized by a unique risk profile, encapsulated by a scalar θ_i. Importantly, θ_i is not a fixed parameter but rather a realization of a random element denoted as Θ_i, reflecting the inherent uncertainty associated with each individual's risk profile.

Continuing from the above, let us use the following to elaborate on the division of observations into homogeneous subpopulations:

Model Assumption 1. *Assume that aside nonhomogeneous random sample $X_1, X_2, \ldots, X_n$, there is some extra information, restated under covariates $Z_1, Z_2, \ldots, Z_m$ where using such extra information, one may partition such random sample into k homogeneous subpopulations, $\mathcal{I}_1, \mathcal{I}_2, \ldots, \mathcal{I}_k$. Moreover, suppose that the risk parameter for such subpopulations can be restated as $\Theta_1, \Theta_2, \ldots, \Theta_k$ where $\mu_l = E_{\Theta_l}(E(X_i^{(l)}|\Theta_l))$, $\sigma_l^2 = E_{\Theta_l}\left[Var(X_i^{(l)}|\Theta_l)\right]$, $\tau_l^2 = Var_{\Theta_l}\left[E(X_i^{(l)}|\Theta_l)\right]$, and $n_l = \#\mathcal{I}_l$ (the notation $\#\mathcal{I}_l$ represents the number of observations or data points in the subpopulation identified as $\mathcal{I}_l$).*

Note 1. *Under the Model Assumption 1, one should note that:*

*(1) The credibility formula for the l*th *subpopulation would be $P_l^{\text{Credibility}} = \alpha_l \bar{X}_l + (1-\alpha_l)\mu_l$, where $\alpha_l = n_l/(n_l + \sigma_l^2/\tau_l^2)$. However, without the above partitioning that the Model Assumption (1) recommended, the credibility formula for the entire of population is $P_{\text{Total}}^{\text{Credibility}} = \alpha\bar{X} + (1-\alpha)\mu$, where $\alpha = n/(n + \sigma^2/\tau^2)$.*

(2) Under the square error loss function, the risk function for the above estimators, respectively, is

$$R_{\text{Total}}^{\text{Credibility}}\left(P_{\text{Total}}^{\text{Credibility}}\right) = E\left[\left(P_{\text{Total}}^{\text{Credibility}} - \mu(\Theta)\right)^2\right] = \frac{1}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}$$

$$R_l^{\text{Credibility}}\left(P_l^{\text{Credibility}}\right) = E\left[\left(P_l^{\text{Credibility}} - \mu(\Theta_l)\right)^2\right] = \frac{1}{\frac{n_l}{\sigma_l^2} + \frac{1}{\tau_l^2}},$$

where $l = 1, 2, \ldots, k$.

Classification of the nonhomogeneous random sample $X_1, X_2, \ldots, X_n$, into k homogeneous subpopulation, suggested by Model Assumption 1, plays a crucial role in insurance pricing, leveraging observable characteristics to group insured individuals with similar expected claims. This classification enables the development of premium rating systems, which express a priori information about new policyholders or insured individuals lacking claims experience. However, a priori classification schemes may not capture all relevant factors for premium rating, as some factors are unmeasurable or unobservable.

Classification of the nonhomogeneous random sample $X_1, X_2, \ldots, X_n$ into k homogeneous subpopulations, as suggested by Model Assumption 1, plays a crucial role in insurance pricing. This process leverages observable characteristics to group insured individuals with similar expected claims. Such classification enables the development of premium rating systems, which provide initial information about new policyholders or insured individuals lacking claims experience. However, these a priori classification schemes might overlook certain relevant factors in premium rating, as some factors are unmeasurable or unobservable.

https://doi.org/10.1017/S1748499525000065 Published online by Cambridge University Press

To address these limitations, a posteriori classification, also known as experience rating, becomes essential. This system re-rates risks by incorporating claims experience into the rating process, resulting in a more equitable and rational price discrimination scheme. By integrating actual claims data, insurers can refine their pricing strategies, better align premiums with individual risk profiles, and promote fairness in insurance pricing.

Diao and Weng (2019) introduced the RTC model, which employs statistical techniques like regression trees to partition the measurable space X into smaller regions where simple models provide accurate fits. In the subsequent step, for each region, they apply the Bühlmann-Straub credibility premium formula to predict the credibility premium. Specifically, given observed data X_i and associated information $Z_{i,1}, \ldots, Z_{i,m}$ for $i = 1, \ldots, n$, a statistical model such as a regression tree determines the probability that the claim experience $X_1, X_2, \ldots, X_n$ arises from Population l. If this probability exceeds 0.5, the credibility premium is predicted using the model developed for Population 1; otherwise, the model developed for Population 2 is used. As the RTC method employs the Bühlmann-Straub credibility premium formula, its credibility premium is given by $P_l^{\text{RTC}} = \alpha_l \bar{X}_l + (1-\alpha_l)\mu_l$ when Population $l = 1, 2, \ldots, k$ is chosen. However, without any classification, the credibility premium is $P_{\text{Total}}^{\text{Credibility}} = \alpha \bar{X} + (1-\alpha)\mu$, where $\alpha = n/(n + \frac{\sigma^2}{\tau^2})$.

The following provides a more general version of Diao and Weng (2019)'s finding for the RTC model under the square error loss function:

Theorem 1. *Under the Model Assumption 1 and the square error loss function:*

(1) The risk function for the Regression Tree Credibility, say, $\underset{\sim}{R}^{\text{RTC}}(\cdots)$*, and the Total Credibility, say,* $R_{\text{Total}}^{\text{Credibility}}(\cdot)$*, respectively, are given by*

$$\underset{\sim}{R}^{\text{RTC}}\left(P_1^{\text{RTC}}, P_2^{\text{RTC}}, \ldots, P_k^{\text{RTC}}\right) = \sum_{l=1}^{k} \omega_l \left[\alpha_l^2 \frac{\sigma_l^2}{n} + (1-\alpha_l)^2 \tau_l^2\right] = \sum_{l=1}^{k} \omega_l \left[\frac{1}{\frac{n_l}{\sigma_l^2} + \frac{1}{\tau_l^2}}\right]$$

$$R_{\text{Total}}^{\text{Credibility}}\left(P_{\text{Total}}^{\text{Credibility}}\right) = \alpha^2 \frac{\sigma^2}{n} + (1-\alpha)^2 \tau^2 = \frac{1}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}$$

where σ_l^2 *represents the variance of the subpopulation* l*,* τ_l^2 *denotes the variance of the subpopulation mean, and* n_l *represents the number of observations in subpopulation* l*.*

(2) Under the extra assumptions $\sum_{l=1}^{k} \omega_l \sigma_l^2 = \sigma^2$ *and* $\sum_{l=1}^{k} \omega_l \tau_l^2 = \tau^2$*, the Regression Tree Credibility premium* P^{RTC}*, dominates the Total Credibility* $P_{\text{Total}}^{\text{Credibility}}$*, that is,* $\underset{\sim}{R}^{\text{RTC}}(\cdots) \leq R_{\text{Total}}^{\text{Credibility}}(\cdot)$*.*

Proof. To drive an induction argumentation, consider the case of $k = 2$. Using the concave function $f(x, y) = \frac{1}{1/x + 1/y}$, we may conclude that

$$\begin{aligned}
\underset{\sim}{R}^{\text{RTC}}\left(P_1^{\text{RTC}}, P_2^{\text{RTC}}\right) &= \sum_{l=1}^{2} \omega_l f\left(\frac{\sigma_l^2}{n}, \tau_l^2\right) \\
&\leq f\left(\omega_1 \frac{\sigma_1^2}{n} + \omega_2 \frac{\sigma_2^2}{n}, \omega_1 \tau_1^2 + \omega_2 \tau_2^2\right) \\
&= R_{\text{Total}}^{\text{Credibility}}(P^{\text{Credibility}}).
\end{aligned}$$

The rest of the proof arrives under an induction argumentation. □

Theorem 1 provides key insights into the risk management implications of utilizing the RTC premium within insurance pricing frameworks. By considering the square error loss function,

https://doi.org/10.1017/S1748499525000065 Published online by Cambridge University Press

it underscores the significance of balancing within-subpopulation variance σ_l^2 with between-subpopulation variance τ_l^2. This balance is essential for optimizing the mixing proportions ω to minimize overall risk, highlighting the importance of strategic decision-making in risk assessment and pricing strategies.

Furthermore, the theorem establishes that incorporating subclassifications based on ω into the MCF premium does not increase the overall risk. In fact, it suggests that the resulting risk is no greater than that associated with P^{Bayes} (Bayes premium) without subclassifications. This implies that subclassification strategies improve risk management efficacy without elevating overall risk levels, emphasizing the effectiveness of the RTC premium in refining pricing strategies and enhancing risk assessment accuracy.

Additionally, Theorem 1 elucidates that partitioning a collective of individual risks does not compromise the prediction accuracy of the RTC formula premium, provided that structural parameters of resulting sub-collectives can be accurately computed. While this theoretical foundation supports the use of partitioning-based premium prediction methods, it also acknowledges the inevitability of statistical estimation errors. The challenge lies in balancing the benefits of partitioning against the adverse effects of estimation errors, necessitating a judicious trade-off in decision-making processes.

Model Assumption 2. *Additional to Model Assumption 1, suppose that for each random variable X_i, for $i = 1, 2, \ldots, n$, there exists additional information $Z_{i,1}, Z_{i,2}, \ldots, Z_{i,m}$ such that using the following logistic regression, one can evaluate the probability that observation belongs to the* lth *subpopulation, that is,*

$$\omega_l = P(X_i \in PoP_l | z_{i,1}, z_{i,2}, \ldots, z_{i,m}) = \frac{1}{1 + \exp\{-\beta_0 - \sum_{j=1}^{m} \beta_l z_{j,l}\}}, \tag{1}$$

for $l = 1, 2, \ldots, k$. Moreover, suppose that the claim experience $X_1, X_2, \ldots, X_n$, given parameter vector $\boldsymbol{\Psi} = (\theta_1, \theta_2, \ldots, \theta_k)'$, follows a k-component normal mixture distribution $\sum_{l=1}^{k} \omega_l N(\theta_l, \sigma_l^2)$, where σ_l^2 are given, and for $l = 1, 2, \ldots, k$, Θ_l has a conjugate prior distribution $N(\mu_l, \tau_l^2)$.

Under Model Assumption 2, Jahanbani *et al.* (2024) introduced the LRC premium as $\sum_{l=1}^{k} \omega_l[\xi_1 \bar{X}_l + (1 - \xi_l)\mu_l]$. They also showed that, under the squared error loss function, its corresponding risk function is:

$$R^{\text{LRC}}(P^{\text{LRC}}) = \sum_{l=1}^{k} \omega_l^2 \left[\xi_l^2 \frac{\sigma_l^2}{n} + (1 - \xi_l)^2 \tau_l^2 \right],$$

where $\xi_l = \sum_{i=0}^{n} \omega_l^i (1 - \omega_l)^{n-i} \binom{n}{i} \frac{i\tau_l^2}{i\tau_l^2 + \sigma_l^2}$.

Also, it has been shown that in the case of $k = 2$, the LRC premium dominates. Jahanbani *et al.* (2024) demonstrated that under certain conditions, the risk function of LRC dominates the RTC premium whenever ω is around 0.5. Additionally, see Jahanbani *et al.* (2022) for practical applications of the RTC model.

3. Mixture credibility formula

This section serves as the main contribution of this article, presenting the introduction of an MCF model. This model integrates the k-means technique into credibility theory, aiming to improve premium prediction accuracy as measured by the risk function under the square error loss function. Through the incorporation of machine learning methods, the MCF model demonstrates enhanced performance compared to traditional approaches in credibility theory.

https://doi.org/10.1017/S1748499525000065 Published online by Cambridge University Press

Our MCF model offers a versatile approach to leverage covariate information for premium prediction, consisting of three key steps. First, we introduce a k-means-based algorithm to utilize covariate information, effectively partitioning a collective of risks into distinct sub-collectives. This segmentation ensures homogeneity within each sub-collective while promoting heterogeneity across sub-collectives in terms of risk profiles. Second, the credibility premium formula is applied to each sub-collective, enabling precise estimation of premiums tailored to the characteristics of each segment. Lastly, we aggregate these segment-specific credibility premiums to derive the overall insurance premium, employing logistic regression to estimate mixture weights for each class. This comprehensive approach maximizes the utilization of covariate information while enhancing the accuracy and flexibility of premium prediction.

The following theorem evaluates the risk function for the MCF when the square error loss function is employed.

Theorem 2. *Under Model Assumption 1, and probabilities of belonging to each class* $(\omega_1, \omega_2, \ldots, \omega_k)$*, the total premium is computed as* $P^{\mathrm{MCF}} = \sum_{l=1}^{k} \omega_l[\alpha_l \bar{X}_l + (1-\alpha_l)\mu_l]$*. Consequently, the total risk function under the square error loss function can be expressed as follows:*

$$R^{\mathrm{MCF}}(P^{\mathrm{MCF}}) = \sum_{l=1}^{k} \omega_l^2 \left[\alpha_l^2 \frac{\sigma_l^2}{n_l} + (1-\alpha_l)^2 \tau_l^2\right].$$

Proof. Using the Model Assumption 1, one may write

$$\begin{aligned}
R^{\mathrm{MCF}}(P^{\mathrm{MCF}}) &= E\left[(P^{\mathrm{MCF}} - \mu(\Theta))^2\right] = E\left[\left(P^{\mathrm{MCF}} - \sum_{l=1}^{k} \omega_l \mu(\Theta_l)\right)^2\right] \\
&= E\left[\left(\sum_{l=1}^{k}\left[\omega_l[\alpha_l \bar{X}_l + (1-\alpha_l)\mu_l] - \omega_l \mu(\Theta_l)\right]\right)^2\right] \\
&= E\left[\left(\sum_{l=1}^{k}\left[\omega_l[\alpha_l \bar{X}_l + (1-\alpha_l)\mu_l] - \omega_l \mu(\Theta_l) \pm \omega_l \alpha_l \mu(\Theta_l)\right]\right)^2\right] \\
&= \sum_{l=1}^{k} \omega_l^2 \alpha_l^2 E\left[\left(\bar{X}_l - \mu(\Theta_l)\right)^2\right] + \sum_{l=1}^{k} \omega_l^2 (1-\alpha_l)^2 E\left[\left(\mu_l - \mu(\Theta_l)\right)^2\right] \\
&= \sum_{l=1}^{k} \omega_l^2 \alpha_l^2 \frac{\sigma_l^2}{n_l} + \sum_{l=1}^{k} \omega_l^2 (1-\alpha_l)^2 Var\left[\mu(\Theta_l)\right] \\
&= \sum_{l=1}^{k} \omega_l^2 \left[\alpha_l^2 \frac{\sigma_l^2}{n_l} + (1-\alpha_l)^2 \tau_l^2\right].
\end{aligned}$$

□

The following theorem demonstrates that the MCF under the square error loss function outperforms the RTC method introduced by Diao and Weng (2019).

Theorem 3. *Under the Model Assumption 1, the MCF under the square error loss function outperforms the RTC, that is,* $R^{\mathrm{MCF}}(\cdot) \underset{\sim}{\leq} R^{\mathrm{RTC}}(\cdots)$.

https://doi.org/10.1017/S1748499525000065 Published online by Cambridge University Press

Proof. Using the result of the above two theorems along with the fact that for all $l = 1, 2, \ldots, k$ $0 \leq \omega_l \leq 1$, one may have:

$$\begin{aligned}
&R^{\mathrm{MCF}}(P^{\mathrm{MCF}}) - \underset{\sim}{R}^{\mathrm{RTC}}\left(P_1^{\mathrm{RTC}}, \ldots, P_k^{\mathrm{RTC}}\right) \\
&\quad = \sum_{l=1}^{k} \omega_l^2 \left[\alpha_l^2 \frac{\sigma_l^2}{n_l} + (1-\alpha_l)^2 \tau_l^2\right] - \sum_{l=1}^{k} \omega_l \left[\alpha_l^2 \frac{\sigma_l^2}{n_l} + (1-\alpha_l)^2 \tau_l^2\right] \\
&\quad = \sum_{l=1}^{k} \left[\alpha_l^2 \frac{\sigma_l^2}{n_l} + (1-\alpha_l)^2 \tau_l^2\right] \omega_l(\omega_l - 1) \leq 0.
\end{aligned}$$

□

We proceed to compare the LRC model with the MCF. Specifically, we demonstrate, under the assumption of normality, that the MCF dominates the LRC model.

Theorem 4. *Under the Model Assumption 2, the MCF premium under the square error loss function dominates the LRC, that is,* $R^{\mathrm{MCF}}(\cdot) \leq R^{\mathrm{LRC}}(\cdot)$.

Proof. In the case of $k = 2$, the difference between the risk functions of two premiums can be restated as

$$\begin{aligned}
&R^{\mathrm{MCF}}(P^{\mathrm{MCF}}) - R^{\mathrm{LRC}}(P^{\mathrm{LRC}}) \\
&\quad = \sum_{l=1}^{2} \omega_l^2 \left[\alpha_l^2 \frac{\sigma_l^2}{n_l} + (1-\alpha_l)^2 \tau_l^2\right] - \sum_{l=1}^{2} \omega_l^2 \left[\xi_l^2 \frac{\sigma_l^2}{n_l} + (1-\xi_l)^2 \tau_l^2\right] \\
&\quad = \sum_{l=1}^{2} \omega_l^2 \frac{\sigma_l^2}{n_l} \left[\alpha_l^2 - \xi_l^2\right] + \sum_{l=1}^{2} \omega_l^2 \tau_l^2 \left[(1-\alpha_l)^2 - (1-\xi_l)^2\right] \\
&\quad = \sum_{l=1}^{2} \omega_l^2 \frac{\sigma_l^2}{n_l} \sum_{i=0}^{n} \omega_l^i (1-\omega_l)^{n-i} \binom{n}{i} \left[\frac{n_l \tau_l^2}{n_l \tau_l^2 + \sigma_l^2} + \frac{i\tau_l^2}{i\tau_l^2 + \sigma_l^2}\right] \left[\frac{n_l \tau_l^2}{n_l \tau_l^2 + \sigma_l^2} - \frac{i\tau_l^2}{i\tau_l^2 + \sigma_l^2}\right] \\
&\qquad + \sum_{l=1}^{2} \omega_l^2 \tau_l^2 \sum_{i=0}^{n} \omega_l^i (1-\omega_l)^{n-i} \binom{n}{i} \left[\frac{\sigma_l^2}{n_l \tau_l^2 + \sigma_l^2} + \frac{\sigma_l^2}{i\tau_l^2 + \sigma_l^2}\right] \left[\frac{\sigma_l^2}{n_l \tau_l^2 + \sigma_l^2} - \frac{\sigma_l^2}{i\tau_l^2 + \sigma_l^2}\right] \\
&\quad = \sum_{l=1}^{2} \omega_l^2 \frac{\sigma_l^2 \tau_l^2}{n_l} \sum_{i=0}^{n} \omega_l^i (1-\omega_l)^{n-i} \binom{n}{i} \tau_l^2 \left[\frac{n_l}{n_l \tau_l^2 + \sigma_l^2} + \frac{i}{i\tau_l^2 + \sigma_l^2}\right] \left[\frac{n_l}{n_l \tau_l^2 + \sigma_l^2} - \frac{i}{i\tau_l^2 + \sigma_l^2}\right] \\
&\qquad + \sum_{l=1}^{2} \omega_l^2 \frac{\sigma_l^2 \tau_l^2}{n_l} \sum_{i=0}^{n} \omega_l^i (1-\omega_l)^{n-i} \binom{n}{i} n_l \sigma_l^2 \left[\frac{1}{n_l \tau_l^2 + \sigma_l^2} + \frac{1}{i\tau_l^2 + \sigma_l^2}\right] \left[\frac{1}{n_l \tau_l^2 + \sigma_l^2} - \frac{1}{i\tau_l^2 + \sigma_l^2}\right] \\
&\quad = \sum_{i=0}^{n} \omega_1^i (1-\omega_1)^{n-i} \binom{n}{i} \omega_1^2 \frac{\sigma_1^2 \tau_1^2}{n_1} \left[\frac{n_1^2 \tau_1^2 + n_1 \sigma_1^2}{(n_1 \tau_1^2 + \sigma_1^2)^2} - \frac{i^2 \tau_1^2 + n_1 \sigma_1^2}{(i\tau_1^2 + \sigma_1^2)^2}\right] \\
&\qquad + \sum_{i=0}^{n} (1-\omega_1)^i (\omega_1)^{n-i} \binom{n}{i} (1-\omega_1)^2 \frac{\sigma_2^2 \tau_2^2}{n_2} \left[\frac{n_2^2 \tau_2^2 + n_2 \sigma_2^2}{(n_2 \tau_2^2 + \sigma_2^2)^2} - \frac{i^2 \tau_2^2 + n_2 \sigma_2^2}{(i\tau_2^2 + \sigma_2^2)^2}\right] \\
&\quad = \sum_{i=0}^{n} \omega_1^i (1-\omega_1)^{n-i} \binom{n}{i} H_{(1,2)}(i),
\end{aligned}$$

https://doi.org/10.1017/S1748499525000065 Published online by Cambridge University Press

where

$$H_{(1,2)}(i) = \omega_1^2 \frac{\sigma_1^2\tau_1^2}{n_1}\left[\frac{n_1^2\tau_1^2 + n_1\sigma_1^2}{(n_1\tau_1^2+\sigma_1^2)^2} - \frac{i^2\tau_1^2 + n_1\sigma_1^2}{(i\tau_1^2+\sigma_1^2)^2}\right] + (1-\omega_1)^2 \frac{\sigma_2^2\tau_2^2}{n_2}\left[\frac{n_2^2\tau_2^2 + n_2\sigma_2^2}{(n_2\tau_2^2+\sigma_2^2)^2} - \frac{(n-i)^2\tau_2^2 + n_2\sigma_2^2}{((n-i)\tau_2^2+\sigma_2^2)^2}\right] \tag{2}$$

which arrives by using the fact that $\binom{n}{i} = \binom{n}{n-i}$.

Now, without loss of generality, assume that the function $H_{(1,2)}(i)$ is a continuous function with respect to i. Therefore, using the first derivative

$$\frac{\partial H_{(1,2)}(i)}{\partial i} = -\omega_1^2 \frac{\sigma_1^2\tau_1^2}{n_1}\frac{2\tau_1^2\sigma_1^2 i}{(i\tau_1^2+\sigma_1^2)^3} + (1-\omega_1)^2\frac{\sigma_2^2\tau_2^2}{n_2}\frac{2\tau_2^2[(n-i)\sigma_2^2 - n_2\sigma_2^2]}{((n-i)\tau_2^2+\sigma_2^2)^3},$$

one may conclude that $\frac{\partial H_{(1,2)}(i)}{\partial i}|_{i=0} > 0$ and $\frac{\partial H_{(1,2)}(i)}{\partial i}|_{i=n} < 0$. On the other hand, since $\frac{\partial^2 H_{(1,2)}(i)}{\partial^2 i} < 0$, one may conclude that $H_{(1,2)}(i)$ as a function of i is a concave function which attains its maximum at $i = n_1 = n - n_2$ in which $H_{(1,2)}(i = n_1) = 0$. Therefore, for the case of $k = 2$, we always have $R^{\mathrm{MCF}}(\cdot) - R^{\mathrm{LRC}}(\cdot) \le 0$.

Now of the general case of $k > 2$, observe that

$$\begin{aligned}\omega_k^i(1-\omega_k)^{n-i} &= (1-\omega_1-\cdots-\omega_{k-1})^i(\omega_1+\cdots+\omega_{k-1})^{n-i}\\ &\le \sum_{l=1}^{k-1}\omega_l^i(1-\omega_l)^{n-i}.\end{aligned}$$

Using the above inequality, the difference between the two risk functions can be bounded above by

$$\begin{aligned}&R^{\mathrm{MCF}}(P^{\mathrm{MCF}}) - R^{\mathrm{LRC}}(P^{\mathrm{LRC}})\\
&= \sum_{l=1}^{k}\omega_l^2\left[\alpha_l^2\frac{\sigma_l^2}{n_l} + (1-\alpha_l)^2\tau_l^2\right] - \sum_{l=1}^{k}\omega_l^2\left[\xi_l^2\frac{\sigma_l^2}{n_l} + (1-\xi_l)^2\tau_l^2\right]\\
&= \sum_{l=1}^{k}\omega_l^2\frac{\sigma_l^2}{n_l}\left[\alpha_l^2 - \xi_l^2\right] + \sum_{l=1}^{2}\omega_l^2\tau_l^2\left[(1-\alpha_l)^2 - (1-\xi_l)^2\right]\\
&= \sum_{l=1}^{k}\omega_l^2\frac{\sigma_l^2}{n_l}\sum_{i=0}^{n}\omega_l^i(1-\omega_l)^{n-i}\binom{n}{i}\left[\frac{n_l\tau_l^2}{n_l\tau_l^2+\sigma_l^2} + \frac{i\tau_l^2}{i\tau_l^2+\sigma_l^2}\right]\left[\frac{n_l\tau_l^2}{n_l\tau_l^2+\sigma_l^2} - \frac{i\tau_l^2}{i\tau_l^2+\sigma_l^2}\right]\\
&\quad + \sum_{l=1}^{k}\omega_l^2\tau_l^2\sum_{i=0}^{n}\omega_l^i(1-\omega_l)^{n-i}\binom{n}{i}\left[\frac{\sigma_l^2}{n_l\tau_l^2+\sigma_l^2} + \frac{\sigma_l^2}{i\tau_l^2+\sigma_l^2}\right]\left[\frac{\sigma_l^2}{n_l\tau_l^2+\sigma_l^2} - \frac{\sigma_l^2}{i\tau_l^2+\sigma_l^2}\right]\\
&= \sum_{l=1}^{k}\omega_l^2\frac{\sigma_l^2\tau_l^2}{n_l}\sum_{i=0}^{n}\omega_l^i(1-\omega_l)^{n-i}\binom{n}{i}\tau_l^2\left[\frac{n_l}{n_l\tau_l^2+\sigma_l^2} + \frac{i}{i\tau_l^2+\sigma_l^2}\right]\left[\frac{n_l}{n_l\tau_l^2+\sigma_l^2} - \frac{i}{i\tau_l^2+\sigma_l^2}\right]\\
&\quad + \sum_{l=1}^{k}\omega_l^2\frac{\sigma_l^2\tau_l^2}{n_l}\sum_{i=0}^{n}\omega_l^i(1-\omega_l)^{n-i}\binom{n}{i}n_l\sigma_l^2\left[\frac{1}{n_l\tau_l^2+\sigma_l^2} + \frac{1}{i\tau_l^2+\sigma_l^2}\right]\left[\frac{1}{n_l\tau_l^2+\sigma_l^2} - \frac{1}{i\tau_l^2+\sigma_l^2}\right]\end{aligned}$$

https://doi.org/10.1017/S1748499525000065 Published online by Cambridge University Press

Table 1. Output of command str(Data)

1	Data.frame	1180255 observations with 6 variables	Values
2	National ID	Character	"0579514676", "3874606295",. . ..
3	Province	Character	"Tehran", "Esfahan", . . .
4	Age	Number	57, 54, 56, 74, 34, 45, 17, ...
5	Gender	Character	"male", "female", "male",. . .
6	Type of insured	Character	"Primary insured", "Subsidiary insured" , . . .
7	Claim size	Real value	9331040, 16900000, 2633500,. . .

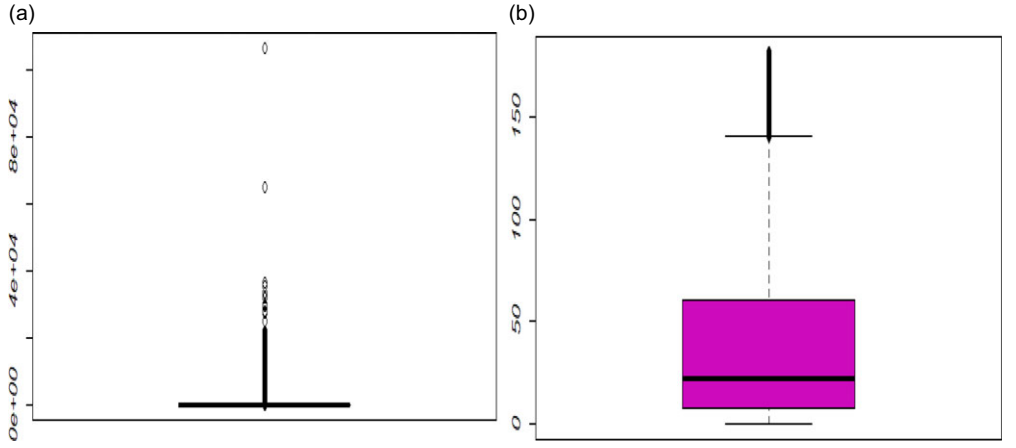


Figure 1. The box plot of damages' size before removing outliers (a) and after removing outliers (b).

$$= \sum_{l=1}^{k} \sum_{i=0}^{n} \omega_l^i (1-\omega_l)^{n-i} \binom{n}{i} \omega_l^2 \frac{\sigma_l^2 \tau_l^2}{n_l} \left[\frac{n_l^2 \tau_l^2 + n_l \sigma_l^2}{(n_l \tau_l^2 + \sigma_l^2)^2} - \frac{i^2 \tau_l^2 + n_l \sigma_l^2}{(i\tau_l^2 + \sigma_l^2)^2} \right]$$

$$\leq \sum_{l=1}^{k-1} \sum_{i=0}^{n} \omega_l^i (1-\omega_l)^{n-i} \binom{n}{i} H_{(l,k)}(i),$$

where $H_{(l,k)}(i)$ is a general version of the $H_{(1,2)}(i)$ given by Equation (2). The same argument, as we did above for the case $k = 2$, leads to desired result for the general case $k > 2$, that is, $R^{\mathrm{MCF}}(\cdot) - R^{\mathrm{LRC}}(\cdot) \leq 0$. □

4. Application to the real-world data

To illustrate the practical application of the above findings, we now consider a real-world dataset from an Iranian insurance company. This dataset contains demographic information (such as gender, age, etc.) as well as the size of claims for 1,180,255 individuals. Table 1 represents the list of available information in the dataset.

Figure 1(a) illustrates the box plot of claim size. However, Figure 1(b) illustrates the box plot of claim size after outlier data has been removed from the dataset. We defined an outlier observation as any observation that exceeds the value of 181,727,944 rials (the Iranian currency), and after removing such outlier data, we obtained 1,179,470 individuals.

Figure 2 shows our attempt to fit a statistical distribution to the claim size.

https://doi.org/10.1017/S1748499525000065 Published online by Cambridge University Press

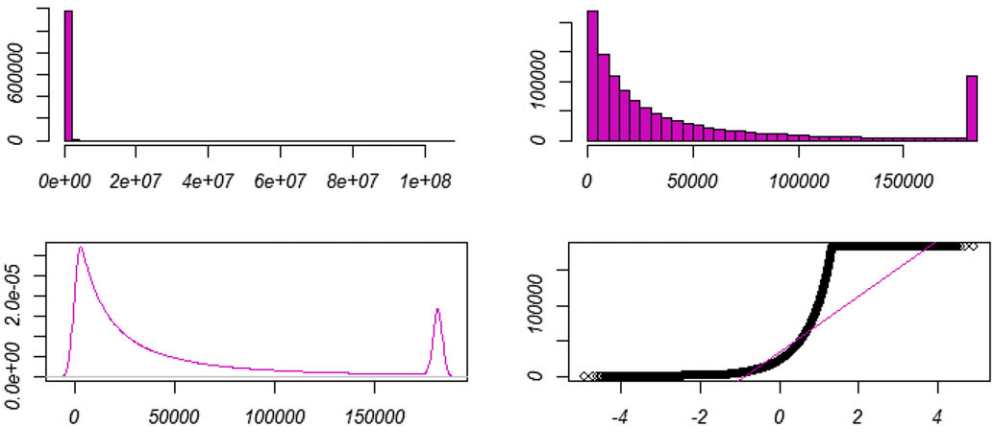


Figure 2. Histogram and density plots of the claim size.

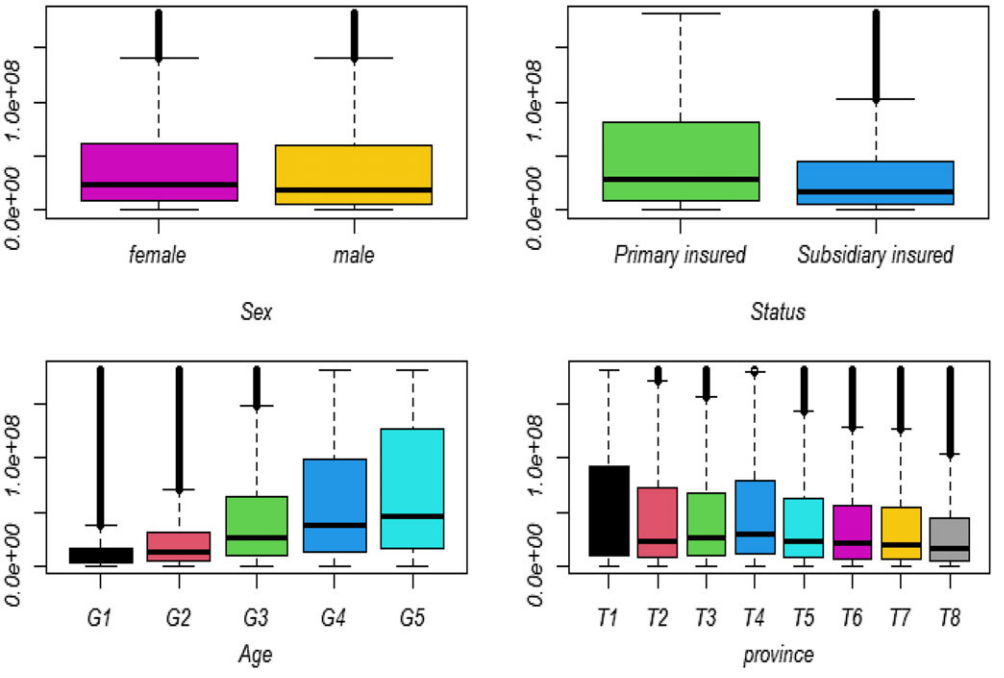


Figure 3. The box plot of damages' size for different categories.

As one may observe, there is considerable non-homogeneity in the data. Therefore, we decided to derive some subpopulations to homogenize the data.

To get started, we categorized the covariates "Age" into 5 classes and "Province" into 8 classes. The Spearman test validated a significant relationship between age, gender, province, type of insured, and claim size. Moreover, the nonparametric Kruskal–Wallis test revealed a significant difference between groups associated with gender, type of insured, age, and province. Figure 3 shows the box plot of claim size regarding the aforementioned covariates.

https://doi.org/10.1017/S1748499525000065 Published online by Cambridge University Press

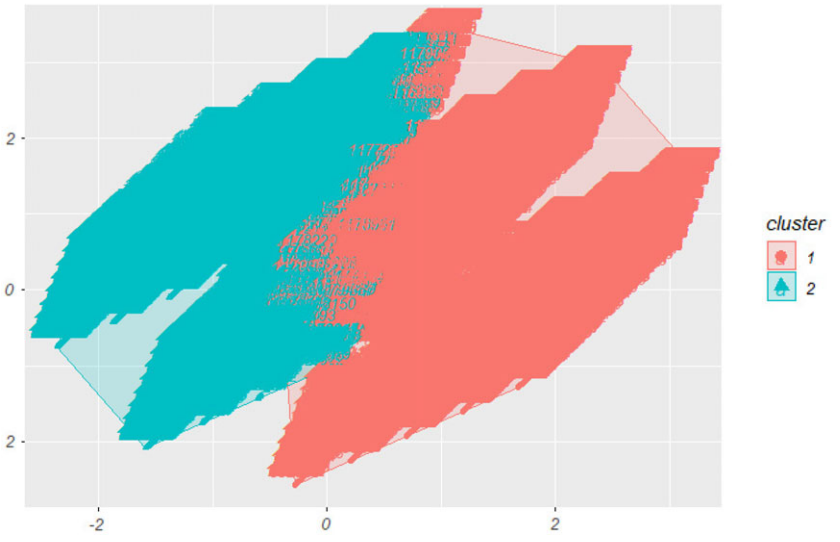


Figure 4. A bivariate visualization of the two subpopulations.

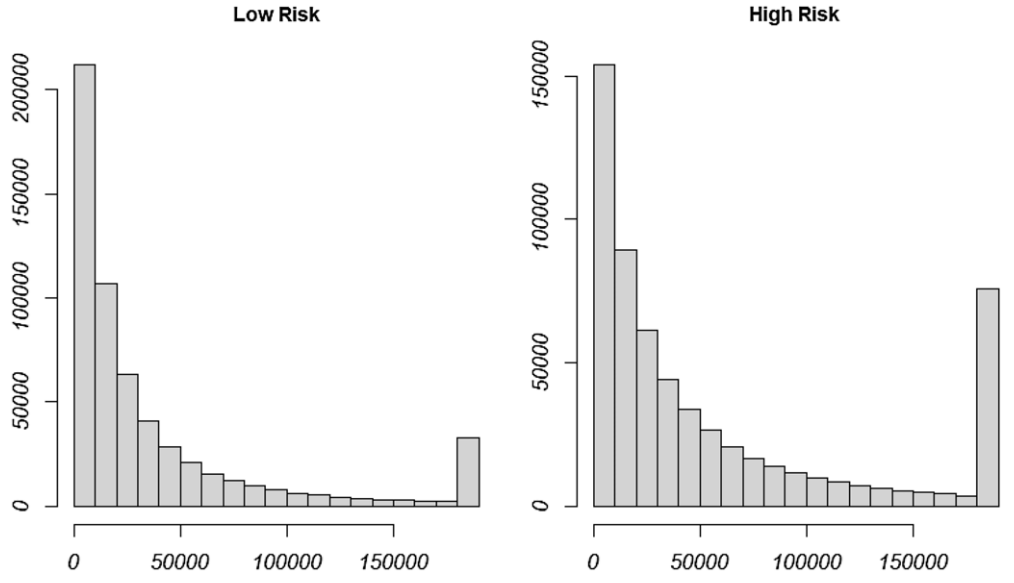


Figure 5. Histogram of claim size for the high-risk and the low-risk classes.

Using Figure 3, one may conclude that these covariates can be employed to define some more homogenous subpopulations.

In the next step, using the *k*-means clustering method, we derived two subpopulations, which classify insured individuals into high-risk and low-risk categories. A bivariate visualization of the two subpopulations is presented in Figure 4. Moreover, Figure 5 illustrates a histogram of claim size for those subpopulations.

https://doi.org/10.1017/S1748499525000065 Published online by Cambridge University Press

Table 2. Logistic regression

Random variable	Estimate	SE	Pr(>|z|)
Intercept	12.88	0.38519	$2e^{-16}$
Province	−0.14	0.03589	$8.87e^{-05}$
Age	9.17	0.19301	$2e^{-16}$
Gender	31.66	0.2564	$2e^{-16}$
Type of insured	−55.8	0.2413	$2e^{-16}$

To determine the probability that a given individual belongs to a given subpopulation, we implement the following logistic regression:

$$\omega_1 = P(Y_i \in Pop_1 | \boldsymbol{Z_i} = \boldsymbol{z_i}) = 1/\{1 + e^{-\beta_0 - \beta_1 z_1 - \beta_2 z_2 - \beta_3 z_3 - \beta_4 z_4}\}, \tag{3}$$

where z_1, z_2, z_3, z_4, respectively, stand for the re-categorized province, age, gender, and type of insured. Using the least square error method against observed data, we estimated the logistic regression parameters. Table 2 reports such estimated parameters.

Now, as an example, consider a 50-year-old single man who lives in a location labeled 6. Moreover, suppose that the logarithms of his 11 years claim experiences are 16.91, 17.68, 15.97, 19.23, 15.63, 16.03, 14.85, 18.41, 14.20, 16.20, and 14.75258.

Using the logistic regression model (3) against his information, with probability 0.2378 ($\omega = 0.2378$), he would fall into Class 1. Therefore, the logarithm of the MCF for his next year is

$$\begin{aligned}
\log(P^{\text{MCF}}) &= \omega \left[\left(\frac{n_1}{n_1 + \frac{2}{\hat{\tau}_1^2}} \right) \bar{x}_1 + \left(1 - \frac{n_1}{n_1 + \frac{\hat{\sigma}_1^2}{\hat{\tau}_1^2}} \right) \hat{\mu}_1 \right] \\
&= +(1-\omega) \left[\left(\frac{n_2}{n_2 + \frac{\hat{\sigma}_2^2}{\hat{\tau}_2^2}} \right) \bar{x}_2 + \left(1 - \frac{n_2}{n_2 + \frac{\hat{\sigma}_2^2}{\hat{\tau}_2^2}} \right) \hat{\mu}_2 \right] \\
&= 0.2378 \left[\left(\frac{597253}{597253 + \frac{2}{1.48}} \right) 17.42 + \left(1 - \frac{597253}{597253 + \frac{2}{1.48}} \right) 15 \right] \\
&\quad +0.7622 \left[\left(\frac{582217}{582217 + \frac{3}{1.51}} \right) 17.83 + \left(1 - \frac{582217}{582217 + \frac{3}{1.51}} \right) 20 \right] \\
&= 17.73251.
\end{aligned}$$

5. Conclusion

The insurance industry relies heavily on accurate predictions of future damages and losses. Various methods, including time series analysis, Bayesian methods, and belief theory, are employed for such predictions. Among these, both classical and Bayesian credibility methods offer accurate and robust predictions.

In this article, under the k-component normal mixture distribution assumption, we introduced the MCF method for insurance premium calculation. This method begins by clustering the insured population into homogeneous subpopulations using data mining techniques, such as k-means. Then, the classical credibility premium is evaluated for each subpopulation. The MCF is a convex combination of those classical credibility premiums. For the case of two subpopulations,

https://doi.org/10.1017/S1748499525000065 Published online by Cambridge University Press

the convex weight can be determined by using logistic regression. We also compared the performance of the MCF method with the RTC method and the LRC method. Our analysis revealed that the MCF method consistently outperforms those methods in terms of the quadratic loss function. This underscores the effectiveness of the MCF method in refining insurance premium calculations and improving risk assessment strategies.

As mentioned earlier, we derive our findings under the k-component normal mixture distribution assumptions; for a possible extension, one may consider how this restrictive assumption can be removed.

Acknowledgments. The authors thank Dr. Shahram Mansour for his valuable comments and suggestions on an earlier version of this manuscript. Moreover, the authors would like to thank the anonymous reviewers for their constructive comments, which improved the theoretical foundation and presentation of this article.

Data availability statement. The data and code that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to restrictions for the insurance company.

Funding statement. There is no financial support for this article.

Competing interests. The authors declare none.

References

Bailey, **A.** (1950). A generalized theory of credibility. *Proceedings of the Casualty Actuarial Society*, **1950**(13), 13–20.

Bühlmann, **H.** (1967). Experience rating and credibility. *ASTIN Bulletin: The Journal of the IAA*, **4**(3), 199–207.

Bühlmann, **H.** (1969). Experience rating and credibility. *ASTIN Bulletin: The Journal of the IAA*, **5**(2), 157–165.

Bühlmann, **H.**, & **Gisler**, **A.** (2005). *A course in credibility theory and its applications*, vol. **317**. Springer.

Bühlmann, **H.** & **Straub**, **E.** (1970). Glaubwürdigkeit für schadensätze. *Bulletin of the Swiss Association of Actuaries*, **70**(1), 111–133.

Cai, **X.**, **Wen**, **L.**, **Wu**, **X.** & **Zhou**, **X.** (2015). Credibility estimation of distribution functions with applications to experience rating in general insurance. *North American Actuarial Journal*, **19**(4), 311–335.

Cheung, **K. C.**, **Yam**, **S. C. P.** & **Zhang**, **Y.** (2022). Satisficing credibility for heterogeneous risks. *European Journal of Operational Research*, **298**(2), 752–768.

Diao, **L.** & **Weng**, **C.** (2019). Regression tree credibility model. *North American Actuarial Journal*, **23**(2), 169–196.

Gómez-Déniz, **E.** & **Vázquez-Polo**, **F. J.** (2022). Exact credibility reference Bayesian premiums. *Insurance: Mathematics and Economics*, **105**, 128–143.

Hickman, **J. C.** & **Heacox**, **L.** (1999). Credibility theory: the cornerstone of actuarial science. *North American Actuarial Journal*, **3**(2), 1–8.

Hong, **L.** & **Martin**, **R.** (2017). A flexible Bayesian nonparametric model for predicting future insurance claims. *North American Actuarial Journal*, **21**(2), 228–241.

Hong, **L.** & **Martin**, **R.** (2018). Dirichlet process mixture models for insurance loss data. *Scandinavian Actuarial Journal*, **2018**(6), 545–554.

Hong, **L.** & **Martin**, **R.** (2022). Imprecise credibility theory. *Annals of Actuarial Science*, **16**(1), 136–150.

Jahanbani, **E.**, **Payandeh Najafabadi**, **A. T.** & **Masoumifard**, **K.** (2022). A credibility approach to predicting healthcare costs: a case study of dental costs. *Iranian Journal of Health*, **5**(2), 96–107.

Jahanbani, **E.**, **Payandeh Najafabadi**, **A. T.** & **Masoumifard**, **K.** (2024). On Bayesian credibility mean for finite mixture distributions. *Annals of Actuarial Science*, **18**(1), 5–29.

Kaas, **R.**, **Goovaerts**, **M.**, **Dhaene**, **J.** & **Denuit**, **M.** (2008). *Modern actuarial risk theory: using R*, vol. **128**. Springer Science & Business Media.

Lau, **J. W.**, **Siu**, **T. K.** & **Yang**, **H.** (2006). On Bayesian mixture credibility. *ASTIN Bulletin: The Journal of the IAA*, **36**(2), 573–588.

Li, **H.**, **Lu**, **Y.** & **Zhu**, **W.** (2021). Dynamic Bayesian ratemaking: a Markov chain approximation approach. *North American Actuarial Journal*, **25**(2), 186–205.

https://doi.org/10.1017/S1748499525000065 Published online by Cambridge University Press

Lo, **A. Y.** (1984). On a class of Bayesian nonparametric estimates: I. *Density estimates. The Annals of Statistics*, **12**(1), 351–357.
Mowbray, **A.** (1914). How extensive a payroll is necessary to give dependable pure premium? *Proceedings of the Casualty Actuarial Society*, **1914**(1), 24–30.
Norberg, **R.** (1979). The credibility approach to experience rating. *Scandinavian Actuarial Journal*, **1979**(4), 181–221.
Payandeh Najafabadi, **A. T.** (2010). A new approach to the credibility formula. *Insurance: Mathematics and Economics*, **46**(2), 334–338.
Payandeh Najafabadi, **A. T. P.**, **Hatami**, **H.** & **Omidi Najafabadi**, **M.** (2012). A maximum-entropy approach to the linear credibility formula. *Insurance: Mathematics and Economics*, **51**(1), 216–221.
Payandeh Najafabadi, **A. T.** & **Sakizadeh**, **M.** (2019). Designing an optimal bonus-malus system using the number of reported claims, steady-state distribution, and mixture claim size distribution. *International Journal of Industrial and Systems Engineering*, **32**(3), 304–331.
Payandeh Najafabadi, **A. T.** & **SakiZadeh**, **M.** (2024). Bayesian estimation for ij-inflated mixture power series distributions using an EM algorithm. *Accepted for publication by: Journal of Thailand Statistician*.
Whitney, **A.** (1918). The theory of experience rating. *Proceedings of the Casualty Actuarial Society*, **1918**(4), 274–292.
Yong, **Y.**, **Zeng**, **P.** & **Zhang**, **Y.** (2024). Credibility theory for variance premium principle. *North American Actuarial Journal*, **29**(1), 1–33. doi: 10.1080/10920277.2023.2299497.
Zhang, **J.**, **Qiu**, **C.** & **Wu**, **X.** (2018). Bayesian ratemaking with common effects modeled by mixture of Polya tree processes. *Insurance: Mathematics and Economics*, **82**, 87–94.

Cite this article: Abed M and Payandeh Najafabadi AT (2026). Mixture credibility formulas, *Annals of Actuarial Science*, **20**, 7–21. https://doi.org/10.1017/S1748499525000065

https://doi.org/10.1017/S1748499525000065 Published online by Cambridge University Press