

Other

Discussion on “*q*-Credibility” by Olivier Le Courtois

Liang Hong^{1a}, Ryan Martin^{2b}

¹ Department of Mathematical Sciences, The University of Texas at Dallas, ² Department of Statistics, North Carolina State University

Keywords: Dirichlet process mixture models, *q*-credibility theory, Bayesian non-parametric models

Variance

Vol. 14, Issue 2, 2021

The classical credibility theory circumvents the challenge of finding the bona fide Bayesian estimate (with respect to the square loss) by restricting attention to the class of *linear* estimators of data. See, for example, Bühlmann and Gisler (2005) and Klugman et al. (2008) for a detailed treatment. Though it is simple to implement and easy to interpret, the classical credibility theory basically only guarantees accurate estimation (i.e., exact credibility) under fairly restricted assumptions such as exponential family models with conjugate priors (Diaconis and Ylvisaker 1979). Therefore, it is natural to seek alternative and more general methods for estimating the mean loss. One such approach is to consider the best *quadratic* estimator of data, as Dr. Le Courtois has done with his *q*-credibility proposal. In this note we provide three comments. The first shows how Le Courtois's Proposition 1.1 can be simultaneously extended and the proof simplified; the second discusses what actuaries can do beyond the classical credibility theory; and the third poses several open problems.

Address for Correspondence: liang.hong@utdallas.edu

1. INTRODUCTION

First, we congratulate Dr. Olivier Le Courtois for this interesting contribution to credibility theory. One of us (Liang) was fortunate enough to meet Dr. Le Courtois and learn about the key results of his paper during the session “Credibility Analysis: Theory, Practice, and Evolution” at the 2016 Joint Statistical Meeting in Chicago, sponsored by both the Society of Actuaries and the American Statistical Association. It is a pleasure to see that the paper is finally in print now.

The classical credibility theory circumvents the challenge of finding the bona fide Bayesian estimate (with respect to the square loss) by restricting attention to the class of *linear* estimators of data. See, for example, Bühlmann and Gisler (2005) and Klugman, Panjer, and Willmot (2008) for a detailed treatment. Though it is simple to implement and easy to interpret, the classical credibility theory basically guarantees accurate estimation (i.e., exact credibility) only under fairly restricted assumptions such as exponential family models with conjugate priors (Diaconis and Ylvisaker 1979). Therefore, it is natural to seek alternative and more general methods for estimating the mean loss.

One such approach is to consider the best *quadratic* estimator of data, as Le Courtois has done with his *q*-credibility proposal.

In this note we provide three comments. The first shows how Le Courtois's Proposition 1.1 can be simultaneously extended and the proof simplified; the second discusses what actuaries can do beyond the classical credibility theory; and the third poses several open problems. To be clear, in no way are any of these comments meant to be critical. Instead, it is our hope that these remarks will stimulate further discussion and developments along this line and ultimately benefit the actuarial science community in general.

2. EXTENDING AND SIMPLIFYING PROPOSITION 1.1

The proof of Proposition 1.1 bears resemblance to its counterpart in the classical theory. That is, the paper applies differential calculus to obtain the solution for a constrained optimization problem. The author is already aware that a short proof exists, and mentions the Hilbert space approach (e.g., Shiu and Sing 2004) in his final remark. While the Hilbert space approach is arguably more elegant, it does

^a Liang Hong is a Fellow of the Society of Actuaries and an Associate Professor in the Department of Mathematical Sciences at The University of Texas at Dallas. He received his PhD in mathematics from Purdue University in 2009. His current research interests are actuarial science, Bayesian analysis & machine learning, foundations of mathematics, and operations research & analytics.

^b Ryan Martin is currently a Professor in the Department of Statistics at North Carolina State University, and his PhD is from the Department of Statistics at Purdue University in 2009. Ryan's research interests include foundations of statistics, generalized Bayes and Gibbs posterior distributions, high- and infinite-dimensional problems, imprecise probability, and mixture models. He is co-author of the monograph *Inferential Models* published by Chapman-Hall/CRC Press in 2016.

put more technical burden on the reader. But an alternative proof, which is both short and elementary, is available too. Indeed, set

$$\begin{cases} Y_i = X_i, & i = 1, \dots, n, \\ Y_i = X_i^2, & i = n+1, \dots, 2n, \\ Y_{2n+1} = X_{n+1} \end{cases}$$

so that the proposed quadratic estimator $\hat{X}_{n+1} = a_{0,q} + \sum_{i=1}^n \alpha_i X_i + \sum_{i=1}^n \beta_i X_i^2$ is transformed to a linear estimator $\hat{Y}_{2n+1} = a_{0,q} + \sum_{i=1}^n \alpha_i Y_i + \sum_{i=n+1}^{2n} \beta_i Y_i$. Then Equations (30) and (32) in Le Courtois (2021) follow immediately the following normal equations in the classical credibility theory (Klugman, Panjer, and Willmot 2008, 583):

$$E(X_{n+1}) = \hat{a}_0 + \sum_{i=1}^n \hat{a}_i E(X_i),$$

$$\text{Cov}(X_k, X_{n+1}) = \sum_{i=1}^n \hat{a}_i \text{Cov}(X_k, X_i), \quad k = 1, \dots, n,$$

where $a_0 + \sum_{i=1}^n a_i X_i$ is the classical credibility estimator of X_{n+1} and $\hat{a}_0, \hat{a}_1, \dots, \hat{a}_n$ are the values of $a_0, a_1, \dots, a_n$ that minimize the corresponding mean squared error. For the remainder of the proof, one can proceed as in Le Courtois (2021).

This technique can also be applied to explore higher-order polynomial functions of the data as estimators. Even more generally, given an appropriately rich dictionary of basis functions $\{\phi_j : j = 1, \dots, J\}$, the estimator

$$\gamma_0 + \sum_{j=1}^J \sum_{i=1}^n \gamma_{ij} \phi_j(X_i)$$

is linear in the data $Y_{ij} = \phi_j(X_i)$. In any case, such transformations are needed if one wants to employ the Hilbert space approach in Shiu and Sing (2004), since the results of their Section 2.2 rely heavily on the assumption that the estimator is linear in data.

3. BEYOND THE CREDIBILITY APPROXIMATIONS

The preceding discussion shows that very complex functions of the data can be considered as linear estimators after a suitable transformation. In this sense, the classical credibility theory essentially covers any basis-type approximation, though the conditions under which exact credibility occurs remain unknown. When Hans Bühlmann first proposed the classical theory in the 1960s, neither the computer technology nor the field of computational statistics was ready for a full Bayesian analysis. With the rapid advances in both fields in the last two to three decades, actuaries are now equipped with the tools needed to build more powerful predictive models beyond what credibility theory has to offer. In particular, actuaries can now seek genuine Bayesian estimates, instead of linear approximations. For example, Hong and Martin (2017) propose a Dirichlet process mixture log-normal model for predicting future insurance claims. When the underlying loss distribution has an arbitrary absolutely continuous density function, this Bayesian nonparametric model is able to produce a full predictive distribution from which the Bayesian premium (i.e.,

predictive mean) and any other feature of interest can be read off.

Example 1. Here we consider a simulated data example similar to Example 20.29 of Klugman, Panjer, and Willmot (2008). That is, claim amounts are assumed to follow the inverse gamma distribution with parameter $\alpha = 4$ and unknown scale parameter θ . The prior distribution for θ is taken to be the gamma distribution with shape parameter 0.5 and scale parameter 100. For our simulation setting, we consider five different values for the true θ , corresponding to the 10th, 25th, 50th, 75th, and 90th percentiles of the gamma prior. For each of these θ values we simulate a sample of size $n = 200$ from the stated inverse gamma and calculate the Bayesian premium $\hat{\mu}_B$, credibility premium $\hat{\mu}_C$, and Dirichlet process mixture premium $\hat{\mu}_{DPM}$. Table 1 gives the estimates and shows that the Dirichlet process mixture premium outperforms the oracle credibility premium. In practice, actuaries will not know the underlying loss distribution, so the credibility premium is subject to further bias from potential model misspecification. However, as Hong and Martin (2018) argue, the Dirichlet process mixture model is more robust and is less susceptible to model misspecification risk.

In addition, the credibility approach provides only a point estimator and ignores other features of the loss distribution. But the Bayesian nonparametric approach allows actuaries to obtain information about these features. For example, Figure 1 shows the plot of the predictive density function along with the histogram of the simulated data for $\theta = 5$. More than that, actuaries can also obtain other interesting quantities such as mean, variance, value-at-risk, and conditional tail expectation; see Hong and Martin (2017) for more details. The R code for this example is available upon request from the first author.

Our conversations with several industry actuaries strike us that it is crucial to have a relatively simple implementation procedure available for the Dirichlet process mixture model so that insurance companies and regulators might be willing to entertain it. Toward that goal, Hong and Martin (2019) propose an easy-to-implement algorithm that does not require the user to know or use Markov chain Monte Carlo methods. It is our hope that in the future practicing actuaries will use this powerful method.

4. SEVERAL OPEN PROBLEMS IN THE CREDIBILITY THEORY

Though powerful alternative methods are now available for obtaining genuine Bayesian estimates, the classical credibility theory remains a cornerstone of actuarial science. There still exist many interesting open problems. Here are a few:

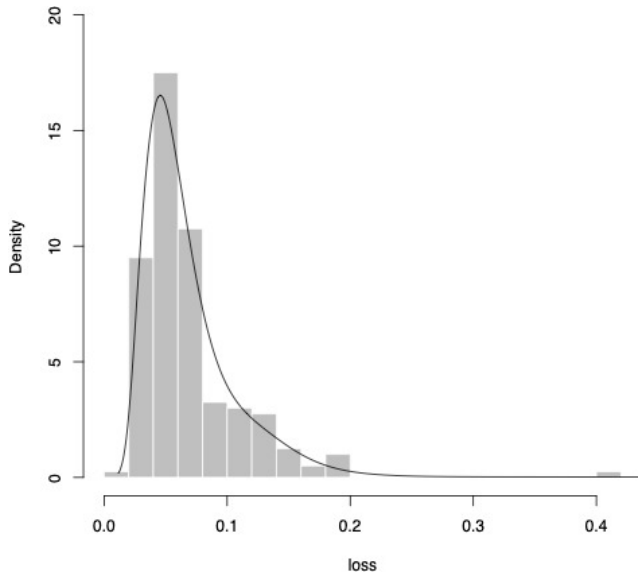
1. Goel (1982) conjectures that if the (posterior) predictive mean is a linear function of the data $X_1, \dots, X_n$, then the marginal distribution of X_k must belong to the exponential family. He imposed the restriction that the distribution function of X_k be of the form

$$\mathcal{F} = \{F_\theta(\cdot) \mid \theta \in \Theta\}, \quad \Theta \subset \mathbb{R},$$

Table 1. Comparison of the Bayesian premium, the oracle credibility premium, and the Dirichlet process mixture premium.

Percentile	θ	$\hat{\mu}_B$	$\hat{\mu}_C$	$\hat{\mu}_{DPM}$
10	0.7895	0.4516	0.5010	0.4480
25	5.0765	0.0702	0.1305	0.0697
50	22.7468	0.0157	0.0775	0.0156
75	66.1652	0.0054	0.0675	0.0054
90	135.2772	0.0026	0.0648	0.0026

The θ values correspond to the stated percentiles of the prior.

**Figure 1. Plot of the histogram of the simulated data with the Dirichlet process mixture estimate of the predictive density overlaid.**

which excludes the Dirichlet process prior as a negative answer. Note that Theorem 1 in Landsman and Makov (1998) does not really give a negative answer to this conjecture because they assume the dispersion parameter $\sigma^2 = \frac{1}{\lambda}$ is known. Hence, their loss distribution is essentially a member in the one-parameter exponential family (also called linear exponential family).

- Goel (1982) also asks whether a linear form of the predictive mean implies that the sample mean $\bar{X}_n$ is a sufficient statistic.
- Regardless of whether the aforementioned conjecture is true, it remains unknown what class of distribu-

tions is implied by a linear predictive mean and how big this family is.

- It is well known that the exponential family with conjugate priors implies exact credibility. Under which conditions will q -credibility be exact too? How about a general n -degree polynomial approximation to the Bayesian estimate?
- A surprising fact in mathematics says that the class of differentiable functions is “dust” in the “universe” of all functions (e.g., Theorem 25.2 in Willard 1970). Can we say the same for the exponential family with conjugate priors relative to arbitrary loss distributions with arbitrary priors?

5. CONCLUSION

We congratulate Dr. Le Courtois again on this interesting extension of the classical credibility theory. Though the classical credibility theory has been investigated for several decades, many interesting open problems remain unsettled. In addition, Bayesian nonparametric models are now available to actuaries for more accurate and efficient prediction. We hope that this discussion stimulates more interest in both directions.

ACKNOWLEDGMENTS

We thank Timothy Wheeler, FCAS, for a fruitful discussion on the challenges academic actuaries must face to make their models more accessible to practicing actuaries and regulators.

Submitted: November 23, 2018 EDT. Accepted: February 15, 2020 EDT. Published: October 28, 2021 EDT.

REFERENCES

- Bühlmann, H., and A. Gisler. 2005. *A Course in Credibility Theory and Its Applications*. New York: Springer.
- Diaconis, Persi, and Donald Ylvisaker. 1979. “Conjugate Priors for Exponential Families.” *The Annals of Statistics* 7 (2): 269–81. <https://doi.org/10.1214/aos/1176344611>.
- Goel, Prem K. 1982. “On Implications of Credible Means Being Exact Bayesian.” *Scandinavian Actuarial Journal* 1982 (1): 41–46. <https://doi.org/10.1080/03461238.1982.10405431>.
- Hong, Liang, and Ryan Martin. 2017. “A Flexible Bayesian Nonparametric Model for Predicting Future Insurance Claims.” *North American Actuarial Journal* 21 (2): 228–41. <https://doi.org/10.1080/10920277.2016.1247720>.
- . 2018. “Dirichlet Process Mixture Models for Insurance Loss Data.” *Scandinavian Actuarial Journal* 2018 (6): 545–54. <https://doi.org/10.1080/03461238.2017.1402086>.
- . 2019. “Online Prediction of Solvency Risk Using Recursive Bayesian Updating.” *Annals of Actuarial Science* 13 (1): 67–79. <https://doi.org/10.1017/s1748499518000039>.
- Klugman, S. A., H. H. Panjer, and G. E. Willmot. 2008. *Loss Models: From Data to Decisions*. 3rd ed. Hoboken: Wiley.
- Landsman, Zinoviy M., and Udi E. Makov. 1998. “Exponential Dispersion Models and Credibility.” *Scandinavian Actuarial Journal* 1998 (1): 89–96. <https://doi.org/10.1080/03461238.1998.10413995>.
- Le Courtois, O. A. 2021. “ q -Credibility.” *Variance* 13:250–64.
- Shiu, E. S. W., and F. Y. Sing. 2004. “Credibility Theory and Geometry.” *Journal of Actuarial Practice* 11:197–216.
- Willard, S. 1970. *General Topology*. Reading, MA: Addison-Wesley.