Principal Applications of Bayesian Methods in Actuarial Science: A Perspective

Udi E. Makov*

Abstract

Bayesian ideas were introduced into actuarial science in the late 1960s in the form of empirical credibility methods for premium setting. The advance of the Bayesian methodology was slow due to its subjective nature and to the computational difficulties associated with the full Bayesian analysis. This paper offers a brief survey of Bayesian solutions to some actuarial problems and discusses the current state of research.

1. Introduction

Bayesian ideas and techniques were introduced into actuarial science in a big way in the late 1960s when the papers of Bühlmann (1967, 1969) and Bühlmann and Straub (1970) laid down the foundation to the empirical Bayes credibility approach, which is still being used extensively in the insurance industry.

Bayesian methodology is used in various areas within actuarial science. This paper does not aim at providing an extensive review of all applications of this methodology, but rather seeks to point out the main areas of application and identify the main themes of contemporary research.

The most important areas of application are:

- Experience rating (including credibility theory), where premiums are set given the accumulated past claims in a portfolio.
- Experience reserving and compound claim modeling. The former deals with the amount of reserves to be kept by an insurance company and the latter with the assessment of the accumulated claims and their impact on the company's financial standing.

An account of Bayesian statistics in actuarial science can be found in a book by Klugman (1992) which concentrates on the Bayesian approach to credibility. For recent review papers see Makov et al. (1996), Schmidt (1998), and Pacakova (1997).

Section 2 introduces the basic ideas of experience rating and credibility theory. It then discusses the broader theoretical and computational issues relating Bayesian solutions to experience rating, which are relevant to actuarial science as a whole. Section 3 discusses experience reserving and compound loss models. Concluding remarks and suggested research directions are given in section 4.

2. Theoretical and Computational Issues: The Case of Experience Rating

In experience rating, θ_{ij} $(i = 1, \ldots I, j = 1, \ldots J)$, the *risk parameter,* is regarded as the total characterization of the risk of contract i at time j. Given θ_{ij}, the actual claims associated with contract $i, X_{i1}, X_{i2} \ldots$ are stochastically independent and follow a distribution $f(x_{ij}|\theta_{ij})$ and the θ's are i.i.d. and follow a prior distribution $U(\cdot)$, commonly called *structure distribution.* The *fair premium* is denoted by $\mu(\theta_{ij}) = \mathrm{E}[X_{ij}|\theta_{ij}]$ and the actual premium is calculated by $\mathrm{E}[X_{i,J+1}|D] = \mathrm{E}[\mu(\theta_{ij})|D]$, where D is the available data, in this context, X_{ij}, $(i = 1, \ldots I,\ j = 1, \ldots J)$. Since $\mathrm{E}[\mu(\theta_{ij})|D]$, the *exact credibility,* was typically hard to calculate, empirical Bayes techniques were employed to produce an estimate of the exact credibility by means of the famous credibility formula $z\bar{x}_i + (1 - z)m$, where m is the mean of $U(\cdot)$ and $\bar{x}_i$ the mean claim of the i^{th} contract;

* Udi E. Makov, Ph.D., is Director, Actuarial Programme, Department of Statistics, University of Haifa, Haifa 31905, Israel, e-mail: makov@rstat.haifa.ac.il.

z, the credibility factor, is chosen to produce the best (m.s.e.) linear empirical Bayes estimator of the exact credibility. Typically $z = aJ/(aJ + s^2)$, where $a = \text{var}[\mu(\theta_{ij})]$, and $s^2 = \text{E}\{\text{var}[X_{ij}|\theta_{ij}]\}$ are unknown and estimated from the data.

The credibility model has been generalized over the years and in most cases maintained its simple linear formula and empirical Bayes flavor. For further reading on the various credibility models, see a brief survey in Makov et al. (1996), Waters (1987), Goovaerts and Hoogstad (1987), Goovaerts et al. (1980), and Herzog (1994).

The empirical Bayes credibility model was a successful practical compromise at a time when opposition to the Bayesian methodology was centered on two major points: (1) opposition to the subjective nature of Bayesian statistics and the search for a more objective tool, and (2) reservations about its applicability, especially as closed-form analytical solutions were not widely available.

The traditional credibility formula constitutes, in a way, a distribution-free methodology. But how accurate is it for particular claim distributions? Herzog (1990) examined the compatibility of the Bayesian and Bühlmann models and showed that the Bühlmann credibility estimate is the best linear approximation of the Bayesian estimate of the fair premium. Following Jewell (1974), it was shown that the credibility estimate is equal to the Bayesian estimate for a large class of problems (exponential family/conjugate priors). See also Schmidt (1980), Goel (1982), and Gerber (1995). Landsman and Makov (1998, 1999a) established that the simple credibility formula is also correct, in a Bayesian sense, for claim distributions belonging to the Exponential Dispersion Model (EDF) (see Jorgensen 1987, 1992), thus allowing computationally simple estimation of the fair premium for members of this family.

Landsman and Makov (1999b, 1999c) established a relationship between *stochastic approximation* and credibility. Using this relationship, a *generalized sequential credibility* was suggested and an optimal stepwise gain sequence derived to produce an estimate of the exact credibility for claim distribution belonging to the location dispersion family and to the symmetric location family. This new methodology allows the derivation of near exact credibility for relevant claim distributions like the log-gamma and log-normal, for which the traditional credibility methods produced only suboptimal solutions. See also Taylor (1977).

The issue of prior specification has recently been dealt with in various ways. Young (1997) suggested using kernel density estimation to estimate the prior distribution of the parameter of interest. To enhance accuracy, Young (2000) suggested employing a loss function, which is a linear combination of a squared-error term and another term designed to reduce divergence. Information measures were used to establish a prior distribution for the dispersion parameter λ of the EDF; Landsman and Makov (1998) used the maximum entropy principle and Landsman and Makov (1999a) minimized the Fisher information. This latter criterion was also used in Landsman and Makov (2001) to establish a prior distribution for λ in conjunction with knowledge on the probability that a claim exceeds a certain threshold, thus allowing for information on tail behavior to affect the premium. Gómez-Déniz et al. (1999) carried out robustness analysis with respect to the prior distribution by considering a contaminated class of prior distribution.

REMARK

All credibility models and their extensions mentioned above (with the exception of Young 2000) assume, in effect, a squared error-loss function and, hence, the choice of the predictive mean as exact credibility. Other loss functions can be considered, each producing a different premium principle. See Heilmann (1989), Kamps (1998), and Gómez-Déniz et al. (1999).

For many years, the implementation of Bayesian models was only possible for simple low dimensional problems (see Klugman, 1992, for analytical approximations suitable for such cases). After all, the evaluation of $\text{E}[X_{i,J+1}|D] = \text{E}[\mu(\theta_{ij})|D]$ would typically require $(I - 1)\ (J - 1) + 1$ multiple integrals, an impossible task for most heterogeneous portfolios. During the last decade, however, there has been an increasing realization that the computations required for full Bayesian analysis can be carried out effectively by means of simulation-based methods and, in particular, Markov chain Monte Carlo (MCMC) methods (Gilks et al., 1996). In effect, the experience-rating problem discussed above can be fully investigated on a PC by means of an hierarchical Bayes model for a portfolio as large as needed. For an introduction to MCMC methods and their actuarial applications see Makov et al. (1996) and Scollnik (1996).

3. LOSS RESERVING AND COMPOUND LOSS MODELS

Loss reserves are needed whenever losses remain unpaid at the end of a year, typically as a result of claims incurred but not yet reported (IBNR) or reported but not yet settled (RBNS).

Let the random variables X_{ij} ($i = 1, \ldots, I; j = 1, \ldots I - i + 1$) denote claim figures (or loss ratios, claim frequencies, and so forth.) for the i^{th} year of origin at the j^{th} development year. The data constitute a triangle (the so called *runoff triangle*), where the upper triangle is given and the lower triangle is to be estimated. Since the problem is crucial for insurance companies, a lot of research has been invested in developing effective methodologies (Taylor 2000). However, relatively little was done from the Bayesian perspective. Verrall (1990) carried a Bayesian analysis of the chain ladder model, which can be interpreted as the two-way model:

$$\log(X_{ij}) = \mu + \alpha_I + \beta_J + \epsilon_{ij},$$

but no attempt was made to treat the unknown variances in a fully Bayesian manner.

A complete hierarchical Bayes model was implemented by Hazan and Makov (2000), where MCMC was used to estimate the parameters of two models: the chain ladder model and a switching regression model that allows the delayed claims to increase up to a point and then decrease over time. MCMC was also employed in Ntzoufras and Dellaportas (n.d.) where various models for outstanding claims problems are discussed and where claim count uncertainty is incorporated. Bayesian methods were also used by Haastrup and Arjas (1996) for estimating claim counts and amounts in individual claim data and by Jewell (1989) and de Alba et al. (1997) for estimating claim counts. State-space models (or Kalman filters), which are dynamic Bayesian models, were also suggested for modeling loss reserves. See De Jong and Zehnwirth (1983).

The benefit of the Bayesian approach is in providing the decision maker with a posterior predictive distribution for every entry in the lower triangle of the runoff triangle. Such a distribution allows the assessment of the required reserves in terms not only of point estimators but also of Bayesian confidence intervals and probabilistic indication on the chance that the amount of a future claim exceeds a given threshold.

The common risk model for aggregate claims $S = Y_1 +, \ldots + Y_N$, where $Y_1, Y_2, \ldots$ represent the amounts of successive claims (assumed i.i.d.), is a dual stochastic process with a claim distribution $f(y|\theta)$ and a count distribution $g(n|\varphi)$. The traditional practice is to fit distributions to N (for various count processes see Schmidt 1998) and to Y and then to employ recursive algorithms to calculate the compound distribution of S. As pointed out by Dickson et al. (1998), the major drawback of this approach is that the fitted distributions are assumed to be known with certainty and, thus, there is no account of the parameter estimation error. For an example of recursive algorithms, see Panjer (1981), Schröter (1991), and Sundt (1992).

The Bayesian model (see, for example, Rytgaard 1990 and Hesselager 1993) is aimed at evaluating the posterior distribution of θ and φ and the predictive distribution of S, given historical data on claims and on the total number of claims at various periods. The MCMC method could allow, in principle, a full implementation of the Bayesian model. No account of a thorough study of this approach has yet been reported. For a preliminary study see Pai (1997).

4. CONCLUSION AND FURTHER WORK

Recent attempts to implement Bayesian methods in actuarial science proved promising. It has long been appreciated by many that the Bayesian paradigm offers a more complete framework of analysis as it allows the uncertainty of the estimation error to be incorporated into the learning process. The posterior distribution contains a wealth of information well beyond the point estimators commonly used in the traditional non-Bayesian analysis.

The MCMC method, which has revolutionized Bayesian statistics, is still to be adopted more extensively in actuarial science. There are new problems to be tackled, and there is a need to adopt the more recent advances in MCMC, which have appeared so far only in the statistical literature (Roberts and Rosenthal 1998).

The reservations about the use of prior distributions can be answered by attempts to implement existing theories relating to noninformative priors and to robust Bayesian analysis in which the sensitivity to the choice of prior distribution is examined (Moreno 2000).

REFERENCES

BÜHLMANN, H. 1967. "Experience Rating and Probability," *ASTIN Bulletin* 4: 199–207.

———. 1969. "Experience Rating and Probability," *ASTIN Bulletin* 5: 157–65.

BÜHLMANN, H., AND E. STRAUB. 1970. "Glaubwurdigkeit fur Schaden-satze," *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathe-matiker* 70: 111–33.

DE ALBA, E., T. M. MORENO, AND M. JUARES. 1997. "Bayesian Estimation of IBNR Reserves." Technical Report, I.T.A.M., Mexico.

DE JONG, P., AND B. ZEHNWIRTH. 1983. "Claims Reserving State Space Models and the Kalman Filter," *Journal of the Institute of Actuaries* 110: 157–81.

DICKSON, D. C. M., L. M. TEDESCO, AND B. ZEHNWORTH. 1998. "Predictive Aggregate Claims Distributions," *The Journal of Risk and Insurance* 65: 689–709.

GERBER, H. U. 1995. "A Teacher's Remark on Exact Credibility," *ASTIN Bulletin* 25: 189–92.

GILKS, W. R., S. RICHARDSON, AND D. J. SPIEGELHALTER eds. 1996. *Markov Chain Monte Carlo in Practice.* London: Chapman & Hall.

GOEL, P. K. 1982. "On Implications of Credible Means Being Exact Bayesian," *Scandinavian Actuarial Journal* 41–6.

GÓMEZ-DÉNIZ, E., A. HERNÁNDEZ-BASTIDA, AND F. J. VÁZQUEZ-POLO. 1999. "The Esscher Premium Principle in Risk Theory: A Bayesian Sensitivity Study," *Insurance: Mathematics & Economics* 25: 387–95.

GOOVAERTS, M. J., AND W. J. HOOGSTAD. 1987. *Credibility Theory.* Rotterdam: National Nederlanden N.V.

GOOVAERTS, M. J., R. KASS, A. VAN HEERWAARDEN, AND T. BAUWELINCKX. 1990. *Effective Actuarial Methods.* North Holland.

HAASTRUP, S., AND E. ARJAS. 1996. "Claims Reserving in Continuous Time; A Nonparametric Bayesian Approach," *ASTIN Bulletin* 26: 139–64.

HAZAN, A., AND U. E. MAKOV. 2000. "Recent Advances in Bayesian Methods in Loss Reserving." Paper presented at the ASTIN Colloq., September, Porto Cervo.

HEILMANN, W. 1989. "Decision Theoretic Foundations of Credibility Theory," *Insurance: Mathematics & Economics* 8: 77–95.

HERZOG, T. N. 1990. "Credibility: The Bayesian Model Versus Bühlmann's Model," *Transactions of the Society of Actuaries* 41: 43–88.

———. 1994. *Introduction to Credibility Theory.* Winsted: ACTEX Publications.

HESSELAGER, O. 1993. "A Class of Conjugate Priors with Applications to Excess-of-loss Reinsurance," *ASTIN Bulletin* 23: 77–93.

JEWELL, W. S. 1974. "Credible Means are Exact Bayesian for Exponential Families," *ASTIN Bulletin* 8: 77–90.

———. 1989. "Predicting IBNYR Events and Delays," *ASTIN Bulletin* 19: 25–55.

JORGENSEN, B. 1987. "Exponential Dispersion Models (with discussion)," *Journal of the Royal Statistical Society Series B* 49: 127–62.

———. 1992. "Exponential Dispersion Models and Extensions: A Review," *International Statistical Review* 60: 5–20.

KAMPS, U. 1998. "On a Class of Premium Principles Including the Esscher Principle," *Scandinavian Actuarial Journal* 71–85.

KLUGMAN, S. A. 1992. *Bayesian Statistics in Actuarial Science.* Boston: Kluwer.

LANDSMAN, Z., AND U. E. MAKOV. 1998. "Exponential Dispersion Models and Credibility," *Scandinavian Actuarial Journal* 89–96.

———. 1999a. "Credibility Evaluation for the Exponential Dispersion Family," *Insurance: Mathematics & Economics* 24: 33–9.

———. 1999b. "On Stochastic Approximation and Credibility," *Scandinavian Actuarial Journal* 1: 15–31.

———. 1999c. "Sequential Credibility Evaluation for Symmetric Location Claim Distributions," *Insurance: Mathematics & Economics* 24: 291–300.

———. 2001. "On Credibility Evaluation and the Tail Area of the Exponential Dispersion Family" *Insurance: Mathematics & Economics* 27: 277–83.

MAKOV, U. E., A. F. M. SMITH, AND Y-H. LIU. 1996. "Bayesian Methods in Actuarial Science," *The Statistician* 45: 503–15.

MORENO, E. 2000. "Global Bayesian Robustness for Some Classes of Prior Distributions," *Robust Bayesian Analysis,* pp. 45–70, Lecture Notes in Statistics. Vol. 152. New York: Springer.

NTZOUFRAS, I., AND P. DELLAPORTAS. (n.d.). "Bayesian Modelling of Outstanding Liabilities Incorporating Claim Count Uncertainty. Submitted for publication.

PACAKOVA, V. 1997. "The Bayesian Inference in Actuarial Science," *Central European Journal of Operations Research and Economics* 5: 255–68.

PAI, J. S. 1997. "Bayesian Analysis of Compound Loss Distributions," *Journal of Econometrics* 79: 129–46.

PANJER, H. H. 1981. "Recursive Evaluation of Family of Compound Distributions," *ASTIN Bulletin* 12: 22–6.

ROBERTS, G. O., AND J. S. ROSENTHAL. 1998. "Markov-Chain Monte Carlo: Some Practical Implications of Theoretical Results," *Canadian Journal of Statistics* 26: 5–31.

RYTGAARD, M. 1990. "Estimation in the Pareto Distribution," *ASTIN Bulletin* 20: 201–16.

SCHMIDT, K. D. 1980. "Convergence of Bayes and Credibility Premium," *ASTIN Bulletin* 20: 167–72.

———. 1998. "Bayesian Models in Actuarial Mathematics," *Mathematical Methods in Operations Research* 48: 117–46.

SCHRÖRER, K. J. 1991. "On a Family of Counting Distributions and Recursions for Related Compound Distributions," *Scandinavian Actuarial Journal* 161–175.

SCOLLNIK, D. P. M. 1996. "An Introduction to Monte Carlo Markov Chain and Their Actuarial Applications," *Casualty Actuarial Society* 158: 114–65.

SUNDT, B. 1992. "On Some Extensions of Panjer's Class of Counting Distributions," *ASTIN Bulletin* 22: 61–80.

TAYLOR, G. 1977. "Abstract Credibility," *Scandinavian Actuarial Journal* 149–68.

———. 2000. *Loss Reserving—An Actuarial Perspective:* Boston: Kluwer.

VERRALL, R. J. 1990. "Bayes and Empirical Bayes Estimation for the Chain Ladder Model," *ASTIN Bulletin* 20: 217–38.

WATERS, H. R. 1987. *An Introduction to Credibility Theory.* London: Institute of Actuaries.

YOUNG, V. R. 1997. "Credibility Using Semiparametric Models," *ASTIN Bulletin* 27: 273–85.

———. 2000. "Credibility Using Semiparametric Models and a Loss Function with a Constancy Penalty," *Insurance: Mathematics & Economics* 26: 151–56.

DISCUSSIONS

MARJORIE A. ROSENBERG*

I would like to commend the author for writing an article in the actuarial literature about the use of Bayesian techniques in actuarial science. Agreeably, Bayesian methods prior to the 1990s were not as visible as they are today due to their computational difficulties of high-level integrals. With the increased power of computers and the advancement of Bayesian statistical techniques, the use of Bayesian methods has increased in the 1990s. Approaches that once seemed too difficult numerically because of complex computations now are commonplace. Simulation of posterior distributions of parameters and of functions of the simulated parameters are easily produced.

The author states that his paper "does not aim at providing an extensive review of all applications of this methodology, but rather seeks to point out the main areas of application and to identify the main themes of contemporary research." He continues by stating that the "most important areas of application are experience rating . . . and experience reserving." These may be the most *prevalent* areas of Bayesian actuarial research due to their historical tradition in actuarial work. While I agree that these are important areas of research, I would not agree that these are the *only* main themes of Bayesian applications in actuarial research.

A couple of years ago, the Society of Actuaries Web site displayed material relating a definition of "actuarial modeling." This material is no longer available on the Web site, but it said: (1) "An actuarial risk is a phenomenon that has economic consequences and is subject to uncertainty with respect to one or more of the actuarial risk variables: occurrence, timing, and severity," and (2) "An actuarial model is a stochastic model, together with a present value model, if applicable, of actuarial risks, based on assumptions about the probabilities that will apply to the actuarial risk variables in the future, including assumptions about the future environment."

Through this definition of actuarial modeling, I view "actuarial research" in a very broad context. Use of Bayesian methods in health care-related research has also increased in the last few years for the same reasons as mentioned above. I would like to introduce readers to some of the literature using Bayesian modeling in the health care practice area. A search on Medline, a health care search engine on the Internet, using the keyword "Bayesian" and covering the years 1966 to the present produced 2,120 citations.[1]

A sample of the journals mentioned in the Bayesian citation include *American Journal of Cardiology, American Journal of Epidemiology American Journal of Public Health, Health Economics, Health Services Research, Inquiry, Journal of Clinical Oncology, Medical Care, Medical*

* Marjorie A. Rosenberg, Ph.D., F.S.A., is a Professor at the University of Wisconsin, School of Business and Department of Biostatistics and Medical Informatics in the Medical School, 975 University Ave., Madison, WI 53706, e-mail: mrosenberg@bus.wisc.edu.

[1] "Medline is produced by the U.S. National Library of Medicine and gathers information from Index Medicus, Index to Dental Literature, and International Nursing, as well as other sources of coverage in the areas of allied health, biological and physical sciences, humanities and information science as they relate to medicine and health care, communication disorders, population biology, and reproductive biology. More than 10.8 million records from more than 3,900 journals are indexed, plus records formally indexed in Healthstar, Bioethicsline and AIDSline. Abstracts are included for about 51% of the records" (from the Medline database description).

Decision Making, Radiology, Statistical Methods in Medical Research, and *Statistics in Medicine.*

I focused on a subset of these journals that contains articles in which actuaries may be interested. A search using a combination of the keyword "Bayesian" and the journal title *Statistics in Medicine* yielded 178 citations. The earliest Bayesian paper listed was by McPherson (1982) concerning a method to incorporate uncertainty into the size of the sample for uncertainties at the design stage of a clinical trial. These 178 articles range in content from clinical trial methodology, results, and economic evaluation to costs, clinical practice decision making, drugs, meta-analysis, surveillance, screening and disease-mapping, and survival models.

As examples, O'Hagan et al. (2001) in a recent article develop a Bayesian computation of the incremental cost-effectiveness acceptability curve for assessing the relative cost-effectiveness of two treatments in health economics, where data on both costs and efficacy are available from a clinical trial. Ashby and Smith (2000) wrote a review article concerning the emergence of evidence-based medicine and the growing use of Bayesian statistics in medical applications. Gelfand and Wang (2000) focused on quantifying the cumulative risk associated with false-positive results on repeated screening procedures for medical conditions, both at the population and the individual level. They developed actuarial models for life table data and added a Cox regression to enable individual level modeling. Daniels and Gatsonis (1997) formulated a hierarchical polytomous regression model and applied it to the analysis of variations in the utilization of alternative cardiac procedures in a national cohort of elderly Medicare patients who had an acute myocardial infarction during 1987. They examined how the rates of cardiac procedures depend on patient-level characteristics, including age, gender, and race, and whether there exist interstate differences and regional patterns in the use of these procedures.

The journal *Health Economics* yielded seven citations for Bayesian studies. One paper by Hamilton (1999) concerned health care costs of Medicare enrollees in health maintenance organizations (HMOs). He found that HMOs select individuals who are less likely to have positive health care expenditures prior to enrollment, but he did not find evidence that HMOs disenroll high-cost patients. Claxton (1999) believed the current regulation of new pharmaceuticals was inefficient because it demanded arbitrary amounts of information, the type of information demanded was not relevant to decision makers, and the same standards of evidence were applied across different technologies. He advocated the use of Bayesian decision theory and an analysis of the value of both perfect and sample information to consider the efficient regulation of new pharmaceuticals.

The journal *Medical Decision Making* yielded 36 citations for Bayesian studies. Parmigiani et al. (1997) developed a framework to quantify uncertainty about costs, effectiveness measures, and marginal cost-effectiveness ratios in complex decision models such as in their example of a stroke-prevention policy. Rosenberg et al. (1999) extended the work of Carlin (1992) to develop a Bayesian method of computing a measure of population health. Observed health-adjusted life expectancy (HALE) is an indicator of population health. HALE is an adjustment to traditional actuarial life expectancy, which weights the year as between 0 and 1 to adjust for quality of life during a year, rather than fixing the measure as 1 for life expectancy. HALE can be used as a tool to help define social policy, as well as a method to assess the needs of a community.[2]

The journal *Medical Care* yielded five citations of Bayesian studies. Cressie's work (1993) involved the prediction of small-area incidence rates from spatially contiguous regions. Miller et al. (1993) estimated what they described as physician "costliness" using outpatient data obtained from a general medicine practice of an urban health care facility. Bay et al. (1983) presented methods used for validating a patient classification system that was based on the concept of types of care (PCTC system). The PCTC system was developed to improve placement decisions for long-term care patients and also to provide

[2] Interestingly, Bayesian methods and graduation has roots in the actuarial literature in the 1960s and later. Jones (1965), Kimeldorf and Jones (1967), and Hickman and Miller (1977) conducted studies of actuarial methods in graduation. Carlin (1992) and Carlin and Klugman (1993) discussed the use of hierarchical Bayesian models in graduation techniques. Graduation methods in actuarial science are mentioned in Makov et al. (1996).

information required for planning in the field of long-term care.

The journal *Statistical Methods in Medical Research* yielded 13 citations of Bayesian studies. Sheiner and Wakefield (1999) discussed the vital role that population (hierarchical) modeling can play in the drug development process, in that population pharmacokinetic/pharmacodynamic models can provide reliable predictions of an individualized dose-exposure-response relationship. "A predictive model of this kind can be used to simulate and hence design clinical trials, find initial dosage regimens satisfying an optimality criterion on the population distribution of responses, and individualized regimens satisfying such a criterion conditional on individual features, such as sex, age, etc." (p. 183).

Clayton (1994) reviewed the work over the last 20 years for the analysis of data with recurrent events, while addressing subject-level heterogeneity. DeAngelis et al. (1993), in their study on AIDS and public health issues were concerned with the uncertainty in the three components of the back calculation method (knowledge of reported AIDS cases, information on the time between HIV infection and onset of AIDS, and assumptions on the rate at which infections occur and the increasingly available information on HIV prevalence) that must be considered to provide realistic projections. Their paper discussed ways of acknowledging uncertainty and suggests a Bayesian formulation.

One of my areas of research is in the statistical application of methods to health care resource utilization and health care policy. With national health expenditures in the United States growing from $73 billion (7% of the gross domestic product) in 1970 to $1.2 trillion (14% of GDP) in 1999 (projected), extensive research is being conducted to model and predict health care spending (see HCFA 1999).

My dissertation focused on the development of a Bayesian statistical model to predict whether a claim was acceptable or not, nonacceptable claims (NACs), from the viewpoint of generally accepted clinical protocols. A statistical system was proposed to monitor NACs (Rosenberg 1999; Rosenberg et al. 1999; and Rosenberg and Griffith 2000) with subsequent research proposing two different approaches to monitor for changes of the rate of nonacceptable claims (Rosenberg 2001a and Rosenberg 2001b). The premise of the statistical system is that inexpensive methods to target audit resources effectively would provide a continuous monitoring of health care costs and help curb NACs.

This research was expanded to examine coding of the diagnosis-related groups (DRGs) for health care claims. Both Medicare and non-Medicare inpatient claims are affected, as private insurers have adopted a payment scheme that is based on the DRG system (Rosenberg et al. 2000). My initial research showed the viability of a statistical system to determine whether an individual claim should be audited to determine whether its DRG coding is incorrect. The statistical system eventually will have three parts: detection, adaptation, and control. At the heart of the statistical system is a hierarchical Bayesian model to predict whether the DRG coding is incorrect.

In conclusion, I wanted to introduce to the actuarial community the breadth of literature in other journals on actuarial-related Bayesian research in health care. Topics span from studies relating to estimating incidence and prevalence rates, meta-analyses of studies, longitudinal studies with respect to modeling costs and decision making in health care utilization.

REFERENCES

ASHBY, D., AND A. F. SMITH. 2000. "Evidence-Based Medicine as Bayesian Decision-Making," *Statistics in Medicine* 19(23): 3291–305.

BAY, K. S., P. LEATT, AND S. M. STINSON. 1983. "Cross-Validation of a Patient Classification Procedure: An Application of the U Method," *Medical Care* 21(1): 31–47.

CARLIN, B. 1992. "A Simple Monte Carlo Approach to Bayesian Graduation," *Transactions of the Society of Actuaries* 44: 55–76.

CARLIN, B., AND S. KLUGMAN. 1993. "Hierarchical Bayesian Whitaker Graduation," *Scandinavian Actuarial Journal* 183–96.

CLAXTON, K. 1999. "Bayesian Approaches to the Value of Information: Implications for the Regulation of New Pharmaceuticals," *Health Economics* 8(3): 269–74.

CLAYTON, D. 1994. "Some Approaches to the Analysis of Recurrent Event Data," *Statistical Methods in Medical Research* 3(3): 244–62.

CRESSIE, N. 1993. "Regional Mapping of Incidence Rates Using Spatial Bayesian Models," *Medical Care,* 31(5 Suppl.): YS60–5.

DANIELS, M., AND C. GATSONIS. 1997. "Hierarchical Polytomous Regression Models with Applications to Health Services Research," *Statistics in Medicine* 16(20): 2311–25.

De Angelis, D., N. E. Day, S. M. Gore, W. R. Gilks, and M. A. McGee. 1993. "AIDS: The Statistical Basis for Public Health," *Statistical Methods in Medical Research* 2(1): 75–91.

Gelfand, A. E., and F. Wang. 2000. "Modelling the Cumulative Risk for a False-Positive Under Repeated Screening Events," *Statistics in Medicine* 19(14): 1865–79.

Hamilton, B. H. 1999. "HMO Selection and Medicare Costs: Bayesian MCMC Estimation of a Robust Panel Data Tobit Model with Survival," *Health Economics* 8(5): 403–14.

Health Care Financing Administration (HCFA). 1999. "National Health Care Expenditures Projections." Online at http://www. hcfa.gov/stats/NHE-Proj/proj1998/proj1998.pdf.

Hickman, J., and R. Miller. 1977. "Notes on Bayesian Graduation," *Transactions of the Society of Actuaries* 29: 1–21.

Jones, D. 1965. "Bayesian Statistics," *Transactions of the Society of Actuaries* 17: 33–57.

Kimeldorf, G., and D. Jones. 1967. "Bayesian Graduation," *Transactions of the Society of Actuaries* 19: 66–112.

Makov, U., A. F. M. Smith, and Y-H. Liu. 1996. "Bayesian Methods in Actuarial Science," *The Statistician* 45(4): 503–15.

McPherson, K. 1982. "On Choosing the Number of Interim Analyses in Clinical Trials," *Statistics in Medicine* 1(1): 25–36.

Miller, M. E., S. L. Hui, W. M. Tierney, and C. J. McDonald. 1993. "Estimating Physician Costliness: An Empirical Bayes Approach," *Medical Care* 31 (5 Suppl.): YS16–28.

O'Hagan, A., J. W. Stevens, and J. Montmartin. 2001. "Bayesian Cost-Effectiveness Analysis from Clinical Trial Data," *Statistics in Medicine* 20(5): 733–53.

Parmigiani, G., G. P. Samsa, M. Ancukiewicz, J. Lipscomb, V. Hasselblad, and D. B. Matchar. 1997. "Assessing Uncertainty in Cost-Effectiveness Analyses: Application to a Complex Decision Model," *Medical Decision Making* 17(4): 390–401.

Rosenberg, M. 1998. "A Statistical Control Model for Utilization Management Programs," *North American Actuarial Journal* 2(2): 77–87.

———. 2001a. "A Statistical Method for Monitoring a Change in the Rate of Nonacceptable Inpatient Claims," *North American Actuarial Journal* 5(4): 74–83.

———. 2001b. "A Decision-Theoretic Method for Assessing a Change in the Rate of Non-Acceptable Inpatient Claims," *Health Services and Outcomes Research Methodology,* 2(1): 19–36.

Rosenberg, M., R. Andrews, and P. Lenk. 1999. "A Hierarchical Bayesian Model for Predicting the Rate of Non-Acceptable In-Patient Hospital Utilization," *Journal of Business and Economic Statistics* 17(1): 1–8.

Rosenberg, M., D. Fryback, and D. Katz. 2000. "A Statistical Model to Detect DRG Upcoding," *Health Services and Outcomes Research Methodology* 1(3–4): 233–52.

Rosenberg, M., D. Fryback, and W. Lawrence. 1999. "Computing Population-Based Estimates of Health-Adjusted Life Expectancy," *Medical Decision Making* 19(1): 90–7.

Rosenberg, M., and J. Griffith. 2000. "A Management Tool for Controlling the Rate of Non-Acceptable Inpatient Hospital Claims," *Inquiry* 36(4): 461–70.

Sheiner, L., and J. Wakefield. 1999. "Population Modelling in Drug Development," *Statistical Methods in Medical Research* 8(3): 183–93.

M. Mendoza*

I would like to congratulate the editors for proposing a paper on Bayesian methods in actuarial sciences for discussion. The author has presented a timely paper that deals with specific applications of the Bayesian paradigm to the solution of some actuarial problems, and I hope that its publication will contribute to further development of this area of research.

To have a general idea of the potential contribution of the Bayesian methods to the actuarial sciences, it might be useful to start with two quotations. From the book by Bernardo and Smith (1994) we have: "Bayesian Statistics offers a rationalist theory of personalistic beliefs in contexts of uncertainty, with the central aim of characterising how an individual should act in order to avoid certain kinds of undesirable behavioral inconsistencies. The theory establishes that expected utility maximisation provides the key to the ways in which beliefs should fit together in the light of changing evidence" (p. 4).

More recently, Lindley (2000) states: "Inference is only of value if it can be used, so the extension to decision analysis, incorporating utility, is related to risk and to the use of statistics in science and law" (p. 293). The point here is that Bayesian statistics is not merely a collection of methods for the analysis of data. It is a theory for making decisions under uncertainty, where both components—probability and utility—are equally important and a full Bayesian analysis includes data analysis, construction of probability models, assessment of the prior information and the utility function, and, finally, making decisions.

In the financial world, making decisions under uncertainty has been recognized to be one of the most important activities. Foccardi and Jones (1998) said: "Risk measurement implies that there is a model of the market that, applied to data, measures risk. But risk management is not limited to the more or less scientific process of measuring risk. Once measured, subjective judg-

* M. Mendoza is a Professor at the Instituto Technologico Autonomo de Mexico, Rio Hondo 1, San Angel, Mexico DF01000, e-mail: mendoza@itam.mx.

ment is used to evaluate and make decisions upon the measurement" (p. 56).

Calculation of an appropriate premium and a reasonable level of reserves, the topics presented in the paper, are financial decision problems that can be formulated within the Bayesian framework in a very general setting. Particular solutions can be obtained corresponding to different choices for both, the probabilistic representation of uncertainty and the specification of the utility function, and even the dependence on these choices might be evaluated by comparing the different particular solutions. Many other actuarial problems can be treated in a similar fashion.

As for the experience-rating problem discussed in the paper, it seems that for some time the proposed solutions had a Bayesian interpretation but they were not proper Bayesian solutions. It is interesting to notice that only recently the situation has been changing. Most of the contributions that explore the assessment of the prior probabilities and the utility function appeared in 1997 or later, and even the more general topic of robustness has been considered. It might be desirable to see in the future many contributions reporting posterior predictive distributions describing the observed experience and proposing utility functions that might be used to calculate premiums corresponding to different strategies and levels of protection.

Even though the author clearly says that the paper does not aim at providing an extensive review of all applications of the Bayesian methodology, it is remarkable the absence of any comment in relation to the topic of Bayesian graduation. This subject has been explored for many years and some recent contributions have appeared that make use of MCMC and other simulation methods. The Bayesian approach to the construction of mortality tables was first considered in relation to the Whittaker graduation procedure, which is not Bayesian but has a Bayesian interpretation. This relationship lead to further study of the conjugate multivariate normal model (see Kimeldorf and Jones 1967 and Hickman and Miller 1977), which apart from some variations of the beta-binomial analysis, was identified for some time as the Bayesian model for graduation.

It is interesting to notice that these models do not incorporate any ordering restriction on the mortality rates and the final product is the posterior distribution for the mortality rates and, more precisely, a Bayes estimate given by the posterior mean. Later, in a paper by Broffitt (1988) an exponential-gamma conjugate model was considered, and different ways of imposing an increasing pattern among mortality rates were explored.

More recently, Dellaportas and Smith (1993) dealt with a closely related problem of making inferences in a class of generalized linear models and proportional hazards models. Carlin (1992) discussed a Monte Carlo approach to Bayesian graduation and, in particular, presented a simulation-based analysis of the model proposed by Broffitt. Dellaportas et al. (2001) concentrated on a Bayesian version of a model originally proposed by Heligman and Pollard (1980), which is intended to give a complete description of the mortality rates across ages, including the so-called accident hump. Mendoza et al. (2001) fitted a linear regression model to a set of transformed death rates and, using predictive arguments, proposed a margin-loaded table, which is mandated for reserving purposes by the current Mexican insurance regulations.

It is worth noticing that, in the usual actuarial practice, interest is focused on the estimation of true mortality rates. Hence, most of the applications of the Bayesian paradigm to the problem of graduation are intended to produce these estimates, which might be used as a mortality (or life) table. In contrast, Mendoza et al. (2001) derived the joint posterior predictive distribution for the mortality rates to be observed in the future for a specific insured population and, on that basis, proposed a strategy to select an appropriate mortality table.

Finally, I would like to congratulate the author for this stimulating paper. As a last comment, and despite the already important number of valuable contributions, I would also have to stress that, in my opinion, the Bayesian approach is far from have being used in its full power to solve even the most well-known actuarial problems, for example, experience rating and graduation. Not only do prior distributions and utility function assessments have to be explored in further detail, but actuaries also have to take advantage of the possibility of producing posterior predictive distributions describing the uncertainty that they will face in the future.

An interesting reference in this direction is Cairns (2000). Other relevant references in the

Bayesian literature that might have an impact in the future application of the Bayesian methods in the actuarial sciences are Bernardo and Muñoz (1993), where Bayesian analysis of demographic data is discussed; Berger and Chen (1993), which makes use of a multinomial model to predict the pattern of a retirement process; Schluter et al. (1997), where a Bayesian procedure is developed for ranking hazardous traffic accident locations; Denison et al. (1998), which deals with the problem of estimating a wide variety of curves using piecewise polynomials, and Walker et al. (1999), where, among other results, a nonparametric approach is proposed to estimate predictive survival curves.

REFERENCES

BERGER, J. O., AND M. CHEN. 1993. "Predicting Retirement Patterns: Prediction for a Multinomial Distribution with Constrained Parameters," *The Statistician* 42: 427–43.

BERNARDO, J. M., AND J. MUÑOZ. 1993. "Bayesian Analysis of Population Evolution," *The Statistician* 42: 541–50.

BERNARDO, J. M., AND A. F. M. SMITH. 1994. *Bayesian Theory.* Chichester, England: Wiley.

BROFFITT, J. D. 1988. "Increasing and Increasing Convex Bayesian Graduation," (with discussion). *Transactions of the Society of Actuaries* 40: 115–48.

CARLIN, B. 1992. "A Simple Monte Carlo Approach to Bayesian Graduation." *Transactions of the Society of Actuaries* 44: 55–76.

CAIRNS, A. J. G. 2000. "A Discussion of Parameter and Model Uncertainty in Insurance," *Insurance: Mathematics & Economics* 27: 313–30.

DELLAPORTAS, P., AND A. F. M. SMITH. 1993. "Bayesian Inference for Generalized Linear and Proportional Hazards Models via Gibbs Sampling," *Applied Statistics* 42(3): 443–59.

DELLAPORTAS, P., A. F. M. SMITH, AND P. STAVROPOULOS. 2001. "Bayesian Analysis of Mortality Data," *Journal of the Royal Statistical Society* (Series A) 164: 275–91.

DENISON, D. G. T., B. K. MALLIK, AND A. F. M. SMITH. 1998. "Automatic Bayesian Curve Fitting," *Journal of the Royal Statistical Society* (Series B) 60: 333–50.

FOCCARDI, S., AND C. JONAS. 1998. *Risk Management: Framework, Methods and Practice.* New Hope, PA: F. J. Fabozzi Associates.

HELIGMAN, L., AND J. H. POLLARD. 1980. "The Age Pattern of Mortality," *Transactions of the Society of Actuaries* 29: 1–21.

HICKMAN, J. C., AND R. B. MILLER. 1977. "Notes on Bayesian Graduation," *Journal of the Institute of Actuaries* 107: 49–80.

KIMELDORF, G. S., AND D. A. JONES. 1967. "Bayesian Graduation," *Transactions of the Society of Actuaries* 19: 66.

LINDLEY, D. V. 2000. "The Philosophy of Statistics," *The Statistician* 49: 293–337.

MENDOZA, M., A. M. MADRIGAL, AND E. GUTIÉRREZ-PEÑA. 2001. "Predictive Mortality Graduation and the Value at Risk: A Bayesian Approach," Technical Report DE-C01.5, Instituto Tecnológico Autónomo de México.

SCHLUTER, P. J., J. J. DEELY, AND A. J. NICHOLSON. 1997. "Ranking and Selecting Motor Vehicle Accident Sites by Using a Hierarchical Bayesian Model," *The Statistician* 46: 293–316.

WALKER, S. G., P. DAMIEN, W. L. PURUSHOTTAM, AND A. F. M. SMITH. 1999. "Bayesian Nonparametric Inference for Random Distributions and Related Functions," *Journal of the Royal Statistical Society* (Series B) 61: 485–527.

FRANCISCO JOSÉ VÁZQUEZ-POLO*

The paper provides a suitable framework for thinking about important aspects of Bayesian actuarial analysis. Most of the methodological and theoretical topics in which Bayesian methodology is developing were discussed, and I congratulate the author. The paper illustrates the utility of Bayesian methods in solving practical issues of credibility theory. Knowing that Bayesian inference can be meaningfully conducted, I will focus my comments on prior sensitivity analysis of the premium principles.

According to robust Bayesian methodology, uncertainty in the prior can be modeled by specifying a class Γ of priors instead of a single one. Bayesian robustness analysis has received substantial attention, and numerous authors have proposed solutions to this problem (Berger 1985, 1990; Berger and O'Hagan 1988; Lavine 1991; Gustafson 1996; Gustafson and Wasserman 1995; Sivaganesan 1988, 1993; and others). An excellent survey of this topic can be found in Berger (1994). However, relatively few papers have been devoted to Bayesian robustness in credibility theory (see Makov 1995 and more recently Gómez et al. 1999, Gómez et al. 2000, and Young 2000).

Robust Bayesian Methodology

Since prior specification is typically imprecise, recent attention focuses on local sensitivity, which measures the effect of perturbations of the inputs on the final answer. Our approach is based on the assumption that the practitioner is unwilling or unable to choose a functional form of the structure function, π_0, but that he may be able to restrict the possible prior to a class that is suitable

* Francisco José Vázquez-Polo is Associate Professor, Department of Quantitative Methods in Economics and Management, University of Las Palmas de Gran Canaria, 35017 Las Palmas de Gran Canaria, Spain, e-mail: polo@empresariales.ulpgc.es.

for quantifying the actuary's uncertainty. Therefore it is of interest to study how the premium for priors in such a class behaves. We use the ϵ-contamination class of priors (see Berger 1985, 1994; Sivaganesan 1991; Sivaganesan and Berger 1989; and Ying-Hsing and Ming-Chung 1997; among others), $\Gamma_\epsilon = \{\pi = (1-\epsilon)\pi_0 + \epsilon q, q \in \mathcal{Q}\}$, where π_0 is the base elicited prior, $\mathcal{Q}$ is the class of allowed contaminations and $\epsilon \in [0, 1]$ measures the uncertainty of the base prior π_0.

The idea of robustness is as follows. Suppose an actuary (decision maker) is looking for a prior distribution of a risk parameter θ, but is unable to choose a single prior. We may say therefore that he is indifferent to the choice of $\pi \in \Gamma_\epsilon$. A natural goal of a robustness study is to find the range of the posterior quantity $P^*_\pi(m)$ when π varies within Γ_ϵ. Thus, attention will be focused on lower and upper bounds for Bayesian premium: $\inf_{\pi\in\Gamma_\epsilon} P^*_\pi(m)$ and $\sup_{\pi\in\Gamma_\epsilon} P^*_\pi(m)$, respectively.

These classes have been used in several situations to measure the sensitivity of quantities that can be expressed in terms of the posterior expectation of parameter functions. Nevertheless, as in other areas of applied statistics, empirical problems raise technical challenges to the state of the art, impinging on the development of new techniques. This is the case of Esscher and variance premium principles. When Esscher or variance premium principles are used, the quantity of interest can be expressed as the ratio of two particular posterior expectations.

In Gómez et al. (1999, 2000), we presented basic results to study the range of the posterior ratio quantities in the form $P_\pi = (\int_\Theta g_1(\theta)\pi(\theta)d\theta)/(\int_\Theta g_2(\theta)\pi(\theta)d\theta)$, as π varies over an ϵ-contamination class

$$\Gamma_\epsilon = \{\pi(\theta) = (1-\epsilon)\pi_0(\theta) + \epsilon q(\theta) | q \in \mathcal{Q}\}, \quad (1)$$

where ϵ reflects the amount of probabilistic uncertainty in a base prior π_0 and $\mathcal{Q}$ is a class of allowable contaminations. For $\mathcal{Q}_1 = \{$All probability distributions$\}$, and $\mathcal{Q}_2 = \{$All unimodal distributions$\}$ we determine the range of P_π as π varies over Γ_ϵ.

The variational problems involved are reduced to finding the extremes of functions of one variable to be solved numerically and the explicit solutions of $\sup_{\pi\in\Gamma} P_\pi(m)$ and $\inf_{\pi\in\Gamma} P_\pi(m)$ to be obtained. A complete development of these techniques can be found in Gómez et al. (1999, 2000), with several illustrations. Among other conclusions, these papers show that unimodality is very convenient, and easy to elicit, for modeling prior information about the risk parameter. It is useful to focus here on the types of actuarial statistical problems in which Bayesian sensitivity analysis can be applied. One future direction presented here is the bimodal structure function.

Bayesian Robustness Analysis in Bimodal Structure Function

Assume that the Bayesian approach is applied to the estimation of net premium. Let Θ be a random variable, and $X_i|\Theta = \theta$, $i = 1, \ldots, t$, the claims or loss amount in subsequent years. We assume that, given θ, the X_i's are conditionally independent and identically distributed random variables.

Suppose that we are interested in studing the situation where the collective has two types of risks; α% are good risks with a low probability of a claim or loss amount, and the other β% are bad risks with a high claim or loss amount probability (see Hewitt 1996; Hewitt and Lefkowitz 1979; and Venter 1991; among others) that can be modeled by two structure functions (prior distributions) $\pi_1(\theta)$ and $\pi_2(\theta)$. Therefore, our prior distribution of θ is given by

$$\pi_0(\theta) = \sum_{i=1}^{2} w_i \pi_i(\theta), \quad (2)$$

with $w_1, w_2 \in \mathbb{R}$, $w_1 + w_2 = 1$.

In experience ratemaking, the actuary takes a claim experience $\mathcal{M} = m$ from the random variables $X_1, X_2, \ldots, X_t$ and uses this information to estimate the unknown fair premium $\mathcal{P}(\theta)$. The posterior distribution of θ, given the likelihood m, is given by

$$\pi_0(\theta|m) = \sum_{i=1}^{2} w_i' \pi_i(\theta|m), \quad (3)$$

where

$$w_i' = \frac{w_i p(m|\pi_i)}{\sum_{i=1}^{2} w_i p(m|\pi_i)}$$

and $p(m|\pi_i) = \int_\Theta f(m|\theta)\pi_i(\theta)d\theta$ is the marginal distribution of $\mathcal{M}$ with respect to prior π_i.

The Bayesian net premium is obtained by

$$\mathcal{P}^*_{\pi_0}(m) = \int_\Theta \left(\int_\chi xf(x|\theta)\, dx \right) \pi_0(\theta|m) d\theta$$
$$= \sum_{i=1}^{2} w_i' \mathcal{P}^*_{\pi_i}(m).$$

Since in our model there are two distinct claim or loss amount generating processes—where some claims or losses are regular and may be described by a p.d.f. $\pi_1(\theta)$, while others are nuisance small claims or losses that may be described by a p.d.f. $\pi_2(\theta)$—our ϵ-contamination class is given by

$$\Gamma^j_\epsilon = \left\{ \pi = (1-\epsilon) \sum_{i=1}^{2} w_i \pi_i + \epsilon \sum_{i=1}^{2} \beta_i q_i,\ q_i \in \mathcal{Q}_i \right\}, \quad (4)$$

with $j = 1, 2$, $\beta_1, \beta_2 \in \mathbb{R}$, $\beta_1 + \beta_2 = 1$.

For $\mathcal{Q}^1_i = \{\text{All probability distributions}\}$, we determine the range of Bayesian net premium as π varies over Γ_ϵ. Now, if we want the model to include distributions with shapes similar to the prior distributions, we can consider the contamination class $\mathcal{Q}^2_i = \{q_i(\theta) : q_i$ is unimodal with the same mode θ_i, as that of $\pi_i\}$.

It is straightforward to rewrite the Bayesian premium under class Γ^1_ϵ, as

$$\mathcal{P}^*_\pi(m) = \frac{(1-\epsilon)\{\sum_{i=1}^2 \alpha_i p(m|\pi_i)\}\mathcal{P}^*_{\pi_0}(m) + \epsilon \int_\Theta g(\theta) f(m|\theta) q(\theta)\, d\theta}{(1-\epsilon)\{\sum_{i=1}^2 \alpha_i p(m|\pi_i)\} + \epsilon \int_\Theta f(m|\theta) q(\theta)\, d\theta}, \quad (5)$$

and

$$\mathcal{P}^*_\pi(m) = \frac{(1-\epsilon)\{\sum_{i=1}^2 \alpha_i p(m|\pi_i)\}\mathcal{P}^*_{\pi_0}(m) + \epsilon \sum_{i=1}^2 \beta_i \int_0^\infty H^{q_i}(z_i)\, dF(z_i)}{(1-\epsilon)\{\sum_{i=1}^2 \alpha_i p(m|\pi_i)\} + \epsilon \sum_{i=1}^2 \beta_i \int_0^\infty H^{i}(z_i)\, dF(z_i)}, \quad (6)$$

under Γ^2_ϵ class. Now, we can easily obtain the ranges for Bayesian premiums using slight modifications of the theorems in Sivaganesan and Berger (1989) and Berger and Moreno (1994). Extensive developments of such results can be found in Gómez et al. (2001).

The Poisson-Gamma Mixture Model

Assume that the number of claims follows a Poisson distribution with parameter θ, while the amount of the individual claim is taken as fixed. Suppose that the prior density of θ is a mixture of two gammas:

$$\pi_0(\theta) = w_1 Gamma(a_1, b_1) + w_2 Gamma(a_2, b_2),$$

where a_1, a_2, b_1, b_2 are positive hyperparameters. Papers using the simple Poisson-gamma are Eichenauer et al. (1988), Klugman (1992), and Freifelder (1974), among others. The following proposition gives lower and upper bounds for Bayesian net premium under the class Γ^1_ϵ. Analogously, proposition 2 gives the bound under the class Γ^2_ϵ. The proofs can be found in Gómez et al. (2001).

Proposition 1

Under the class Γ^1_ϵ, the lower (upper) bound for the Bayesian net premium is given by

$$\inf_{\theta\in\Theta} (\sup) \frac{\mathcal{R}_1 \mathcal{P}^*_{\pi_0}(m) + \mathcal{R}_2(\theta)}{\mathcal{R}_1 + \mathcal{R}_3(\theta)},$$

where:

$$\mathcal{R}_1 = (1-\epsilon) \sum_{i=1}^{2} \alpha_i \frac{a_i^{b_i}}{(b_i-1)!} \frac{(b_i+tm-1)!}{(a_i+t)^{b_i+tm}},$$

$$\mathcal{R}_2(\theta) = \epsilon \theta^{tm+1} e^{-tm},$$

$$\mathcal{R}_3(\theta) = \mathcal{R}_2(\theta)/\theta,$$

and

$$\mathcal{P}^*_{\pi_0}(m) = \sum_{i=1}^{2} \alpha_i' \frac{b_i + tm}{a_i + t}.$$

Proposition 2

Under the class $\pi \in \Gamma_\epsilon^2$, the lower (upper) bound for the Bayesian net premium is given by $\inf_{z_1, z_2 \geq 0}(\sup)\mathcal{R}(z_1, z_2)$, being

$$\mathcal{R}(z_1, z_2) = \frac{\mathcal{R}_1 \mathcal{P}^*_{\pi_0}(m) + \sum_{i=1}^{2} (1/z_i)\beta_i \int_{\theta_i}^{\theta_i+z_i} \mathcal{R}_2(\theta)d\theta}{\mathcal{R}_1 + \sum_{i=1}^{2} (1/z_i)\beta_i \int_{\theta_i}^{\theta_i+z_i} \mathcal{R}_3(\theta)d\theta} \quad \text{if } z_1, z_2 > 0,$$

$$\mathcal{R}(z_1, 0) = \frac{\mathcal{R}_1 \mathcal{P}^*_{\pi_0}(m) + (1/z_1)\beta_1 \int_{\theta_1}^{\theta_1+z_1} \mathcal{R}_2(\theta)d\theta + \beta_2\mathcal{R}_2(\theta_2)}{\mathcal{R}_1 + (1/z_1)\beta_1 \int_{\theta_1}^{\theta_1+z_1} \mathcal{R}_3(\theta)d\theta + \beta_2\mathcal{R}_3(\theta_2)} \quad \text{if } z_1 > 0,$$

$$\mathcal{R}(0, z_2) = \frac{\mathcal{R}_1 \mathcal{P}^*_{\pi_0}(m) + \beta_1\mathcal{R}_2(\theta_2) + (1/z_2)\beta_2 \int_{\theta_2}^{\theta_2+z_2} \mathcal{R}_2(\theta)d\theta}{\mathcal{R}_1 + \beta_1\mathcal{R}_3(\theta_1) + (1/z_2)\beta_2 \int_{\theta_2}^{\theta_2+z_2} \mathcal{R}_3(\theta)d\theta} \quad \text{if } z_2 > 0,$$

and

$$\mathcal{R}(0, 0) = \frac{\mathcal{R}_1 \mathcal{P}^*_{\pi_0}(m) + \sum_{i=1}^{2} \beta_i\mathcal{R}_2(\theta_i)}{\mathcal{R}_3(\theta) + \sum_{i=1}^{2} \beta_i\mathcal{R}_3(\theta_i)},$$

where $\mathcal{R}_1$, $\mathcal{R}_2(\theta)$, $\mathcal{R}_3(\theta)$, and $\mathcal{P}^*_{\pi_0}(m)$ are as in Proposition 1.

To illustrate the above ideas, note the following numerical illustration. We shall use $\beta_i = w_i$, $i = 1, 2$. Also, we have included a measure that does not depend on the premium measurement units. This is the relative sensitivity (RS) index (Sivaganesan 1991):

$$RS^j = \frac{1}{2\mathcal{P}^*_{\pi_0}(m)} \left[\sup_{\pi \in \Gamma_\epsilon^j} \mathcal{P}^*_{\pi_0}(m) - \inf_{\pi \in \Gamma_\epsilon^j} \mathcal{P}^*_{\pi_0}(m)\right] \times 100\%, \ (j = 1, 2).$$

EXAMPLE

Let $X|\theta$ have a Poisson distribution with parameter θ and $\pi_0(\theta) = 0.8\text{Ga}(2, 4) + 0.2\text{Ga}(3, 30)$. With this elicitation, the actuary knows that the two modal values are around 1.5 and 10 (i.e., $\theta_1 = 1.5$ and $\theta_2 = 10$), and that claims above 5 are less frequent than smaller claims.

Let $m = 4$, and $m = 12$. Table 1 contains the standard Bayesian premium. This particular situation corresponds to $\epsilon = 0$, that is, no errors in the process of elicitation. The bounds on the Bayesian net premium are given in Figure 1 for the classes Γ_ϵ^1 and Γ_ϵ^2. Figure 2 displays their RS factor.

Table 1
Standard Poisson-Gamma Model

m	$\mathcal{P}^*_{\pi_1}(m)$	$\mathcal{P}^*_{\pi_2}(m)$	α'_1	α'_1	$\mathcal{P}^*_{\pi}(m) = \sum_{i=1}^{2} \alpha'_i \mathcal{P}^*_{\pi_i}(m)$
4	3.666	5.384	0.998	0.002	3.670
12	10.333	11.538	0.001	0.999	11.538

Concluding Remarks

A basic assumption of credibility theory is that the values of the parameters of the probability distribution of loss are unknown. In this case, the premium that the company charges is the Bayesian premium. This premium requires that the decision maker, the actuary, can define a probability distribution for the values of the unknown parameters of this loss process, the prior distribution.

Nevertheless, there will clearly be many prior distributions other than π_0 which are also compatible and, hence, could be used in place of π_0.

Figure 1
Ranges of Bayesian Net Premium: Poisson-Gamma Mixture Model.

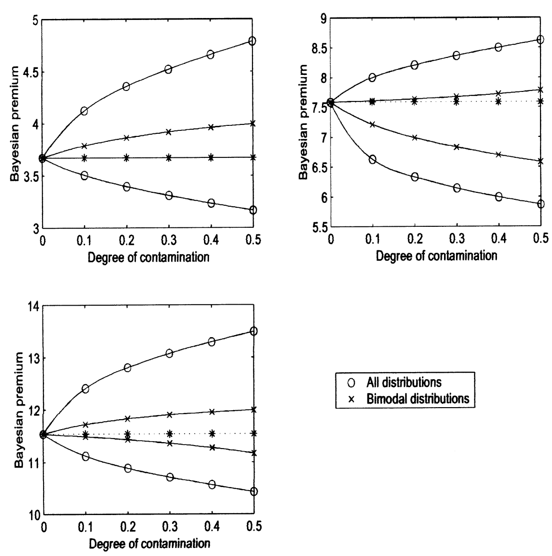
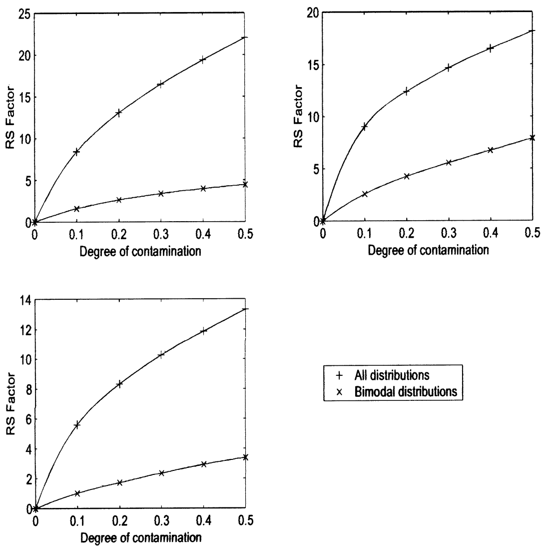


Figure 2
RS Factor: Poisson-Gamma Mixture Model.



Although bimodal models are common in actuarial practice, there has been relatively little research from a Bayesian point of view.

This discussion has presented a simple approach to develop bimodal Bayesian sensitivity analysis. The mixture of two unimodal densities appears appropriate when the researcher agrees that a bimodal structure function is necessary. This is justified in our model, where prior π_0 is given by a convex sum of two prior distributions, π_1 and π_2. This leads to the question of Bayesian robustness, which has been treated in this discussion using the ϵ-contamination class. We have seen that bimodality effects are very important in modeling subjective beliefs about risk parameter. Finally, all theorems and results presented here can be used for alternative premium calculation principles (Heilmann 1989) such as exponential, Esscher, and variance, among others.

ACKNOWLEDGMENT

Research partially supported by the DGUI del Gob. Aut. de Canarias under grant PI2000/061, Canary Islands, Spain.

REFERENCES

BERGER, J. O. 1985. *Statistical Decision Theory and Bayesian Analysis.* New York: Springer.

———. 1990. "Robust Bayesian Analysis: Sensitivity to the Prior," *Journal of Statistical Planning and Inference* 25: 303–28.

———. 1994. "An Overview of Robust Bayesian Analysis" (with discussion), *Test* 3: 5–124.

BERGER, J. O., AND E. MORENO. 1994. "Bayesian Robustness in Bidimensional Models: Prior Independence," *Journal of Statistical Planning and Inference* 40: 161–76.

BERGER, J. O., AND A. O'HAGAN. 1988. "Range of Posterior Probabilities for the Class of Unimodal Priors with Specified Quantiles." In *Bayesian Statistics 3,* J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, eds. New York: Oxford University Press.

EICHENAUER, J., J. LEHN, AND S. RETTIG. 1988. "A Gamma-Minimax Result in Credibility Theory," *Insurance: Mathematics and Economics* 2: 49–57.

FREIFELDER, L. 1974. *Statistical Decision Theory and Credibility Theory Procedures,* proceedings of the Berkeley Actuarial Research Conference on Credibility, Credibility Theory, and Applications. Berkeley: Academic Press.

GÓMEZ, E., A. HERNÁNDEZ, AND F. VÁZQUEZ-POLO. 1999. "The Escher Premium Principle in Risk Theory: A Bayesian Sensitivity Study," *Insurance: Mathematics and Economics* 25: 387–95.

———. 2000. "Robust Bayesian Premium Principles in Actuarial Science," *Journal of the Royal Statistical Society* (Series D, The Statistician) 49(2): 241–52.

GÓMEZ, E., A. HERNÁNDEZ, J. M. PÉREZ, AND F. VÁZQUEZ-POLO. 2001. "Premium Sensitivity Using Bimodal Prior Distributions," Poster presented at SAMO, Madrid, Spain.

GUSTAFSON, P. 1996. "Local Sensitivity of Inferences to Prior Marginals," *Journal of the American Statistical Association* 91: 774–81.

GUSTAFSON, P., AND L. WASSERMAN. 1995. "Local Sensitivity Diagnostic for Bayesian Inference," *The Annals of Statistics* 24(6): 2153–67.

HEILMANN, W. 1989. "Decision Theoretic Foundations of Credibility Theory," *Insurance: Mathematics and Economics* 8: 77–95.

HEWITT, C. 1966. "Distribution by Size of Risk" (with discussion), *Proceedings of the Casualty Actuarial Society* 53: 106–17.

HEWITT, C., AND B. LEFKOWITZ. 1979. "Methods for Fitting Distributions to Insurance Loss Data," *Proceedings of the Casualty Actuarial Society* 66: 139–60.

KLUGMAN, S. 1992. *Bayesian Statistics in Actuarial Science.* Boston: Kluwer Academic Publisher.

LAVINE, M. 1991. "Sensitivity in Bayesian Statistics: The Prior and the Likelihood," *Journal of the American Statistical Society* 86: 400–3.

MAKOV, U. 1995. "Loss Robustness Via Fisher-Weighted Squared-Error Loss Function," *Insurance: Mathematics and Economics* 16: 1–6.

SIVAGANESAN, S. 1988. "Range of Posterior Measures for Priors with Arbitrary Contaminations," *Communications in Statistics and Theory Methods* 17(5): 1591–612.

———. 1991. "Sensitivity of Some Posterior Summaries When the Prior is Unimodal with Specified Quantiles," *The Canadian Journal of Statistics* 19(1): 57–65.

———. 1993. "Robust Bayesian Diagnostic," *Journal of Statistical Planning and Inference* 35: 171–88.

SIVAGANESAN, S., AND J. O. BERGER. 1989. "Ranges of Posterior Measures for Priors with Unimodal Contaminations," *The Annals of Statistics* 17(2): 868–89.

VENTER, G. 1991. "Effects of Variations from Gamma-Poisson Assumptions," *Proceedings of the Casualty Actuarial Society* 78: 41–56.

YING-HSING, L., AND Y. MING-CHUNG. 1997. "Posterior Robustness in Simultaneous Estimation Problem with Exchangeable Contaminated Priors," *Journal of Statistical Planning and Inference* 65: 129–43.

YOUNG, V. R. 2000. "Credibility Using Semiparametric Models and a Loss Function with a Constancy Penalty" *Insurance: Mathematics and Economics* 26(1): 151–6.

JAMES C. HICKMAN* AND DONALD A. JONES†

We are grateful to the author for providing a perspective on current activity in applying Bayesian methods in actuarial science. The author's enthusiasm for Markov chain Monte Carlo (MCMC) methods for implementing Bayesian methods is shared by all of us who are impressed by the coherence of Bayesian methods but have been frustrated by the daunting amount of integration required to obtain results from many practical models.

The purpose of this discussion is to add historical background to the contemporary view of the author. We will review the birth of Bayesian statistics and the crucial role played by an actuary in this event. Next will come sections on the early role of Bayesian methods in graduation and credibility. These two topics illustrate, in a natural fashion, how Bayesian methods provide a framework for combining new information with existing beliefs. The discussion will end with the suggestion that Bayesian time series analysis can become an important tool for actuaries, especially in the design and management of health, pension, and social insurance systems.

* James C. Hickman is Emeritus Professor and Dean at the University of Wisconsin, Grainger Hall, 975 University Ave., Madison, WI 53706.

† Donald A. Jones is a Professor at Oregon State University, Department of Mathematics, 368 Kidder Hall, Corvallis, OR 97331-4605, e-mail: jonesd@math.orst.edu.

Creation

Richard Price (1723–91) was a remarkable gentleman of the age of enlightenment. By profession he was a dissenting minister but as a scholar he had an impact on philosophy, statistics, public finance, demography, political science, and actuarial science. From 1768 to 1791, Price served as a consultant to the Society for Equitable Assurances on Lives and Survivorships. His first assignment was to calculate the probability of survival of a complicated status involving two women and one man of known ages. He constructed the Northampton Life Table and wrote *Observations on Reversionary Payments: On Schemes for Providing Annuities for Widows and for Persons of Old Age; On the Method of Calculating Values of Assurances on Lives; and On the National Debt* (Price 1771). This two-volume publication seems to have left no subject untouched. Benjamin Franklin was, however, sufficiently impressed to call it "the foremost production of human understanding that this century afforded us." The book went through seven editions and was the principal actuarial textbook until well into the 19th century. The story of Price's contribution to actuarial science and Franklin's good opinion of Price's book can be found in Ogborn (1962, chap. 7).

With the passage of about 240 years, we know little of the partnership of Price and Thomas Bayes in examining the fundamental issues in statistics. Bayes was also a dissenting minister and was probably about 20 years older than Price. Price mentions an argument developed by Bayes in his 1758 essay on morals, "A Review of the Principal Questions and Difficulties in Morals," but there is little evidence of earlier joint work. It is possible that both men had studied mathematics with John Eames, a friend of Isaac Newton, at different times in a dissenting academy. Price's essay on morals and a critical review of his philosophy can be found in Hudson (1970).

Bayes, who died in 1761, left £100 and his scientific papers to Price. In his will, Bayes referred to his legatee as "Richard Price, now I suppose, a preacher at Newington Green." Bayes died only four months after executing his will. His address for Price was correct. These events are described by Stigler (1986, chap. 3).

Price decided that the most promising paper left by Bayes was an essay, "Toward Solving a Problem in the Doctrine of Chance." It was read by Price to the Royal Society on Dec. 23, 1763, more than two and one-half years after Bayes' death. Bayes had been elected a Fellow of the Royal Society (FRS) in 1742, despite an apparent absence of any technical publications. Price was elected FRS in 1765 because of his role in presenting Bayes' paper after adding an introduction and appendix. In the appendix, Price worked on the problem of evaluating the incomplete beta function. There are other aspects of the paper that have been attributed to Price, but this issue will remain forever shrouded.

The paper had almost no immediate impact on science. About 20 years later it was rediscovered in connection with Laplace's work on the same problem. Yet today what are called "Bayesian" methods are applied in most scientific fields, and the field of statistics is divided into camps called the Bayesians and the "frequentists."

The consequence of Bayes' work was a mathematical procedure by which an investigator can change existing (a priori) beliefs about a proposition in light of new evidence (data) to produce a revised (posterior) estimate of the degree of belief in the proposition. Although the original beliefs of two investigators might differ, if they both adopted Bayesian methods and processed the same new evidence, their posterior views of the proposition under review would move closer together. Ultimately Bayes' ideas prompted a reevaluation of the fundamental scientific concepts of evidence and causality.

Bayes' paper is difficult for those of our generation to read because, like Newton, he adopted a geometric mode of exposition. Yet to Price, and others of the age of enlightenment, this was the standard way for science to proceed. The analysis showed a clear understanding of the philosophical principles on which it was built, and it contained some subtle mathematics.

Despite Price's critical role, the resulting school of thought is called Bayesian and not Prician, yet Price was present at its creation and contributed to its propagation.

Graduation

In recent decades, graduation lost its position as a leading topic in actuarial education and research. Yet to gain a balanced perspective on Bayesian methods in actuarial science, we must return to graduation. Just as Richard Price played a critical role at the creation, E.T. Whittaker's role in Bayesian methods of graduation is central. Whittaker's (1920) address to the Faculty of Actuaries on the topic "On Some Disputed Questions of Probability" generated a spirited discussion. These questions centered on the interpretation of probability. Some of the questions remain disputed and Whittaker's insights remain valuable.

Whittaker's graduation method involved the minimization of the loss function $L = F + hS$. In this function, F is a measure of the lack of fit, S a measure of lack of smoothness, and h a positive constant. Whittaker's ideas have motivated generations of research-minded actuaries and other applied mathematicians. To many students, it has been presented as a mathematical programming problem with the positive constant h under the control of the investigator. But Whittaker provided a remarkable and complete Bayesian motivation of L, and the constant h is the ratio of two variances. The Bayesian development is perhaps most easily found in Whittaker and Robinson (1944).

Later Kimeldorf and Jones (1967) extended Whittaker's model to include a prior distribution that incorporates a vector of prior means and a prior covariance matrix. A stream of modifications to Bayesian graduation methods followed. One of these contributions deserves special notice. Carlin (1992) reported on Bayesian graduation and in performing what would have been tedious approximate integrations, he used the Gibbs sampler, a Monte Carlo integration method that belongs to the MCMC class. We believe this was among the first actuarial applications of MCMC.

Credibility

The author correctly identifies Bühlmann's (1967) credibility paper as the beginning of a flood of Bayesian-based research in actuarial science. Yet approximately a quarter of a century before 1967, there was startling Bayesian-based research on credibility. In a sequence of papers published between 1942 and 1950, Arthur Bailey developed the rudiments of a Bayesian foundation for credibility. Bailey's (1950) ultimate paper summarized his ideas under the encompassing title "Credibility Procedures: Laplace's Generalization of Bayes' Rule and the Combination of Collateral Knowledge." In this paper, he presents cogent criticisms of existing statistical methods, quotes Price's introduction to Bayes' paper and discusses Laplace's extension. He also applies these ideas to credibility. Most of the credibility ideas developed in the 1960s and 1970s have visible roots here. For example, the use of least squares regression lines to approximate posterior means is a thread running through much of the analysis. The fact that these least squares are exact for several common examples is illustrated. All of this was put forth over 20 years before Ericson (1970) and Jewell (1974) established that, when a conjugate prior distribution (the prior and posterior distributions are of the same type) is used with a likelihood (data distribution) that is a member of the linear exponential family of distributions, then the least squares linear approximation to the posterior mean is exact.

Because he was a self-taught Bayesian, and many of our present day concepts and notation were in the future, Bailey is hard to read. Nevertheless, his work demonstrates that Bayesian ideas were an important ingredient of credibility theory before 1967.

Time Series

We recommend Bayesian time series analysis, especially to actuaries working with health, pension, and social insurance systems. The analysis of time series data of health costs, wage rates, and the size of the labor force is required in these applications. The conventional advice is to analyze the data but combine this analysis with collateral information and prior knowledge. Bayesian time series analysis provides a disciplined procedure for this mixing of information. Rosenberg and Young (1999) provide a primer on Bayesian time series analysis. They use MCMC to complete the analysis and illustrate the methods with unemployment data.

REFERENCES

BAILEY, A. L. 1942. "Sampling Theory in Casualty Insurance," *Proceedings of the Casualty Actuarial Society* 29(59): 50–93.

———. 1945. "A Generalized Theory of Credibility," *Proceedings of the Casualty Actuarial Society* 32(62): 13–20.

———. 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(67): 7–23.

———. 1950. "Discussion," *Journal of the American Teachers of Insurance* (now the *Journal of Risk and Insurance*) 17: 17–24.

CARLIN, B. P. 1992. "A Simple Monte Carlo Approach to Bayesian Graduation." *Transactions of the Society of Actuaries* 44: 55–76.

ERICSON, W. A. 1970. "On the Posterior Mean and Variance of a Population Mean," *Journal of the American Statistical Association* 65(330): 649–52.

JEWELL, W. S. 1974. "Credible Means are Exact Bayesian for Exponential Families," *ASTIN Bulletin* 8: 77–90.

HUDSON, W. D. 1970. *Reason and Right: A Critical Examination of Richard Price's Moral Philosophy.* London: Macmillan.

KIMELDORF, G. S., AND D. A. JONES. 1967. "Bayesian Graduation," *Transactions of the Society of Actuaries* 19(1): 66–127.

OGBORN, M. E. 1962. *Equitable Assurances: The Story of Life Assurance in the Experience of the Equitable Life Assurance Society, 1762–1962.* London: Allen and Unwin.

PRICE, R. 1771. *Observations on Reversionary Payments: On Schemes for Providing Annuities for Widows and for Persons in Old Age; On the Method of Calculating the Values of Assurance on Lives; and On the National Debt.* London: T. Cadell and W. Davis.

ROSENBERG, M. A., AND V. R. YOUNG. 1999. "A Bayesian Approach to Understanding Time Series Data," *North American Actuarial Journal* 3(2): 130–43

STIGLER, S. M. 1986. *The History of Statistics.* Cambridge, MA: The Belknap Press.

WHITTAKER, E. T. 1920. "On Some Disputed Questions of Probability," *Transactions of the Faculty of Actuaries* 8: 163–206.

WHITTAKER, E. T., AND G. ROBINSON. 1944. *The Calculus of Observations*. 4th ed. London and Glasgow: Blackie and Son.

ENRIQUE DE ALBA*

I would first like to thank the author for his stimulating paper that deals with two of the main applications of Bayesian methods in actuarial science. Since Bayesian methods can essentially be applied whenever there is a statistical inference problem, there has always been a Bayesian option in actuarial science. Its infrequent use in the past

* Enrique de Alba is a Professor at ITAM, Division de Actuaria, Estadistica y Matematicas, Rio Hondo No. 1, Mexico, D.F. 01000, Mexico, and at the University of Waterloo, Department of Statistics and Actuarial Science, 200 University Ave. West, Waterloo, ON N2L 3G1, Canada, e-mail: dealba@itam.mx.

has had to do with the general situation existing in statistics over the last decades as well as with the fact that its application has been, until recently, computationally cumbersome or downright impossible.

In fact, I believe that the Bayesian contents of Bühlmann (1967, 1969) have largely been minimized. Some authors only talk about the "structure" distributions and use Bayesian results to derive formulas for estimation of credibility factors. Once the formulas are obtained, the Bayesian framework is largely forgotten. Credibility factors are then estimated by empirical Bayes (EB) or other methods. Available information can and should be used to specify the prior distribution. This is different from using the data to estimate the hyperparameters or some features of the prior distribution, or the decision rule. This is what EB does.

I would like to make some observations about EB methods that I believe are not generally taken into account when using them, specifically in actuarial science. The EB formulation generally used in credibility falls under the *parametric empirical Bayes* label of Morris (1983), where the prior distribution of the risk parameter θ, say $f(\theta|\phi)$, is assumed to be in some parametric class with unknown hyperparameters ϕ. In a fully Bayesian approach these would be known, perhaps elicited from available prior information. Alternatively, a hierarchical model may be used and a distribution specified for ϕ. However, in EB these hyperparameters are typically estimated by some classical (non-Bayesian) method and substituted into the prior distribution, assuming they are the known fixed values (see Berger 1985).

The analysis then proceeds using the prior $f(\theta|\hat{\phi})$. This, however, ignores the fact that the hyperparameters were estimated and the errors introduced by this are not considered in any of the conclusions. Klugman (1992) mentions that none of the existing credibility analyses allows for the extra variability. Hence, I do not totally agree with the author that EB was a successful compromise, since its true meaning and properties usually are not fully taken into account. Furthermore, EB methods have the problem that their optimality properties are essentially asymptotic. Yet, their small sample properties, if known, are usually not considered.

In addition, they essentially provide point estimates. In general, there are no interval estimates for the credibility premiums. When they are obtained they are usually "naïve" intervals (see Carlin and Louis 1996), where the extra variability due to estimation of the parameters is not incorporated. Theoretically, this will yield intervals that are too short, and adjustments may be necessary. Klugman (1992) does provide an example where he shows that the difference is negligible. The use of full Bayesian hierarchical models automatically accounts for the added uncertainty. Hence, strictly speaking, EB is not Bayesian because it does not admit a distribution for ϕ, the hyperparameters of the prior distribution $f(\theta|\phi)$ of the risk parameter (O'Hagan 1994).

There are two points of Bayesian methods that I believe are not stressed sufficiently in this paper. One is that actuarial science is a field where very frequently one has considerable prior information, be it in the form of global or industry-wide information (experience) or in the form of tables. In this respect, it is indeed surprising that Bayesian methods have not been used more intensively up to now. This availability of information would make one question the author's statement that they have not been used more intensely because of their subjective nature. There is a wealth of "objective" prior information available to the actuary.

Another advantage of Bayesian methods that I feel is not emphasized sufficiently in the paper is the possibility of obtaining complete posterior or predictive distributions. Actuarial science is a field in which adequate understanding and knowledge of the complete distribution is essential. In addition to expected values, we are looking at certain characteristics of probability distributions, for example, probability of ruin, value at risk (VAR), and so on. Some of the usual approximations currently in use ignore the possible skewness of the resulting distributions. It is automatically incorporated when using Bayesian methods, whether analytically or numerically.

One advantage of full Bayesian methods is that the posterior distribution for the parameters is essentially the exact distribution; that is, it is true given the specific data used in its derivation. They automatically account for all the uncertainty in the parameters.

I agree with the author that one cause of the low use of Bayesian methods has been the fact that closed analytical forms had not been available and numerical approaches had been too cumbersome to carry out. However, things have changed drastically in recent years. The availability of software that allows one to obtain the posterior or predictive distributions by direct Monte Carlo methods or by Markov chain Monte Carlo (MCMC) has opened a broad area of opportunities for the applications of these methods in actuarial science. He mentions that no attempt has been made to use Bayesian models to evaluate the posterior distribution of the parameters in the frequency and severity distributions of the claims. In addition to the reference given in the paper, this is done in de Alba (2001) by using direct Monte Carlo methods.

One area of application of Bayesian methods in actuarial science that has totally been left out of the paper is *graduation.* This topic has received a great deal of attention from actuaries, including Bayesians; see Kimeldorf and Jones (1967) and Hickman and Miller (1977). Indeed, London (1985), in a literature review of this topic, describes the use of Bayesian methods and provides examples. He deals explicitly with Bayesian graduation and indicates how prior information may be used from standard published tables. London mentions the case where

> . . . [W]e needed mortality rates for making pension calculations for a large group of employees, but had no data derived from the recent experience of this group. We would be likely to choose rates from a published table which was based on the experience of another group, with characteristics as similar to our group as possible. If then, a few years later, we did a study of mortality in our group, and chose to graduate it . . . , these standard table rates would be logical candidates for the prior mean vector . . . (1985, p. 79).

He is clearly outlining the "updating" or learning process that emerges naturally in Bayesian methods. He later points out that a statistical model for mortality can be summarized in the form of a probability distribution, with any number of "distinct" mortality tables obtained from it. A procedure along these lines has been applied in Mendoza et. al. (2001), who use a Bayesian predictive approach to graduation that incorporates a desired level of protection against deviations in mortality. It can be interpreted in terms of VAR.

Among a host of other applications, Scollnik (2001) illustrates how the Kimeldorf-Jones (Kimeldorf and Jones 1967) graduation model can be implemented via MCMC. The main advantage here over the traditional analysis is that, in addition to means, percentiles associated with the posterior distribution can easily be determined.

Another area of growing interest in actuarial science is the use of generalized linear models (Habermann and Renshaw 1996). Specifically, these authors consider survival modeling and claims reserving, among other applications. These applications can actually be carried out from the Bayesian point of view, as has been done in Verrall (1990) and de Alba (2001) in the case of claims reserving. Mendoza et. al. (2001) apply them in their graduation models.

Thus, in addition to the principal applications of Bayesian methods in actuarial science mentioned in the paper, there many potential uses of these procedures that can contribute to a better understanding of actuarial problems. Actuarial science can use them to exploit the wealth of available prior information in a formal way.

REFERENCES

BERGER, J. O. 1985. *Statistical Decisions and Bayesian Analysis.* 2nd ed. New York: Springer-Verlag.

BÜHLMANN, H. 1967. "Experience Rating And Probability," *ASTIN Bulletin* 4: 199–207

———. 1969. "Experience Rating And Probability," *ASTIN Bulletin* 5: 157–65

CARLIN, B. P., AND T. A. LOUIS. 1996. *Bayes and Empirical Bayes Methods for Data Analysis,* New York: Chapman & Hall.

DE ALBA, E., AND M. JUAREZ. 2001. "Bayesian Estimation of Outstanding Claims Reserves," *Research Report 01-01,* Institute of Insurance and Pension Research, University of Waterloo.

HABERMAN, S., AND A. E. RENSHAW. 1996. "Generalized Linear Models and Actuarial Science," *The Statistician* 45(4): 407–36.

HICKMAN, J. C., AND R. B. MILLER. 1977. "Notes on Bayesian Graduation," *Transactions of the Society of Actuaries* 29: 1–21.

KIMELDORF, G. S., AND D. A. JONES. 1967. "Bayesian Graduation," *Transactions of the Society of Actuaries* 19, 66-112.

KLUGMAN, S. A. 1992. *Bayesian Statistics in Actuarial Science.* Boston: Kluwer.

LONDON, R. L. 1985. "*Graduation: The Revision of Estimates.*" Winsted, CT: ACTEX Publications.

MENDOZA, M., A. M. MADRIGAL, AND E. GUTIÉERREZ-PEÑA. 2001. "Predictive Mortality Graduation and the Value at Risk: A

Bayesian Approach." Working Paper DE-C01.5, ITAM, Mexico.
MORRIS, C. 1983. "Parametric Empirical Bayes Inference: Theory and Applications," *Journal of the American Statistical Association* 78: 47–65.
O'HAGAN, A. 1994. *Kendall's Advanced Theory of Statistics, Vol. 2B, Bayesian Statistics.* London: Edward Arnold.
SCOLLNIK, D. P. M. (2001) "Actuarial Modeling with MCMC and BUGS," *North American Actuarial Journal* 5(2): 96–124.
VERRALL, R. J. 1990. "Bayes and Empirical Bayes Estimation for the Chain Ladder Model," *ASTIN Bulletin* 20: 217–38.

DAVID P. M. SCOLLNIK*

I'd like to thank the author for his summary review of Bayesian methods appearing in actuarial science. I would also like to take this opportunity to draw attention to two recent papers absent from this summary.

The first of these is by Promislow and Young (2000) and relates to the discussion of experience rating in Section 2. This paper continues previous work of these authors relating to credibility estimators developed in accordance with a principle of equity. They make the case that a person who should be charged 1 unit but is actually charged 10 units is being treated more unfairly than a person who should be charged 1,001 units but is actually charged 1,010 units.

This situation can arise when credibility premiums are developed using a loss function like the squared error, which emphasizes the absolute difference between the charged premium and the true premium. Their solution is to develop equitable credibility premiums using an entropy loss function, instead, so that a measure of the relative difference between the charged premium and the true premium is minimized in place of the usual squared error. They develop simple credibility formulas and demonstrate that these are exact for some interesting cases when the claim distribution belongs to the linear exponential family.

The second paper is by Scollnik (2001) and relates to the discussion in Sections 3 and 4. This paper describes how a number of different actuarial models, including models of the sort described in Klugman (1992); Makov, Smith, and Liu (1996); Pai (1997); and Rytgaard (1990), can be implemented and analyzed in accordance with the Bayesian paradigm using Markov chain Monte Carlo (MCMC) via the BUGS (Bayesian inference Using Gibbs Sampling) suite of software packages. BUGS is a specialized software package for implementing MCMC-based analyses of full probability models in which all unknowns are treated as random variables. Its programming language is easily understood and allows the user to make a straightforward specification of the full probability model under consideration. The Windows version of BUGS is known as WinBUGS, and it provides a graphical interface to the BUGS language. Scollnik (2001) also includes many additional references to recent papers in actuarial science incorporating Bayesian methods.

* David P.M. Scollnik, A.S.A., Ph.D., is Associate Professor, Department of Mathematics and Statistics, University of Calgary, Calgary, Alberta, Canada, T2N 1N4, e-mail: scollnik@ucalgary.ca.

REFERENCES

KLUGMAN, S. A. 1992. *Bayesian Statistics in Actuarial Science.* Boston: Kluwer.
MAKOV, U. E., A. F. M. SMITH, AND Y-H. LIU. 1996. "Bayesian Methods in Actuarial Science," *The Statistician* 45: 503–15.
PAI, J. S. 1997. "Bayesian Analysis of Compound Loss Distribution," *Journal of Econometrics* 79: 129–46.
PROMISLOW, S. D., AND V. R. YOUNG. 2000. "Equity and Exact Credibility," *ASTIN Bulletin* 30(1): 3–11.
RYTGAARD, M. 1990. "Estimations in the Pareto Distribution," *ASTIN Bulletin* 20: 167–72.
SCOLLNIK, D. P. M. 2001. "Actuarial Modeling with MCMC and BUGS," *North American Actuarial Journal,* 5(2): 96–124.

Additional discussions on this paper can be submitted until April 1, 2002. The author reserves the right to reply to any discussion. Please see the Submission Guidelines for Authors on the inside back cover for instructions on the submission of discussions.

AUTHOR'S REPLY

I would like to thank the discussants for their valuable comments and additional references, and to commend the editors of the *North American Actuarial Journal* (*NAAJ*) for their decision to turn my contribution into a discussion paper. The end result is an in-depth review of the important role of Bayesian statistics in actuarial science. I share the hope expressed by Mendoza that it will contribute to a further development of this field of research. After all, as de Alba put it, "Bayesian methods can essentially be applied whenever there is a statistical inference problem.”

When I singled out the principal applications of Bayesian methods in actuarial science, my aim was to point out some areas in which Bayesian methods have been applied extensively (credibility and loss reserving). In so doing, I could not cover all areas of significant importance. Thanks to the discussants, additional areas and new directions for future research are now exposed to the readers of the *NAAJ*.

Hickman and Jones placed the Bayesian inroads into the actuarial profession in a fascinating historical perspective and rightly emphasized the pioneering contribution of Bailey. Rosenberg focused on Bayesian modeling in the important area of health care. She scanned major journals and exposed current fields of research. While many of the thousands of articles that emerged from her search of the word “Bayesian” in Medline are not directly related to actuarial practice, the many recent publications she quoted bear witness to a lively research area reported in journals unfamiliar to many actuaries.

Another area of application is that of graduation, which has attracted considerable Bayesian attention over the years. Hickman and Jones, de Alba, and Mendoza outlined milestones of Bayesian graduation from the Bayesian interpretation of Whittaker graduation to Markov chain Monte Carlo (MCMC) solutions.

Mendoza took a wider look at the Bayesian approach by supplementing probabilistic modeling with utility functions as part of the broader decision theory context. Consequently, the end product of the analysis is not merely the evaluation of posterior distributions (or estimation of the model’s parameters), but the taking of a decision that maximizes the expected utility. This approach requires both the assessment of prior distributions and utility functions, an important area for future research. Mendoza’s suggestion that actuaries have to take advantage of the posterior predictive distribution should attract serious consideration. This is also emphasized by de Alba.

Vázquez-Polo provided us with a lucid introduction to Bayesian prior robustness. He further developed Bayesian robustness analysis in bimodal structure function and obtained interesting results, which are suitable whenever the heterogeneity of the portfolio (and that of the structure parameter) can be expressed by means of a mixture model. In so doing, he laid out the direction future research is likely to follow.

Hickman and Jones recommend using Bayesian time series analysis in the fields of health, pension, and social insurance systems. De Alba reports the growing interest in the use of Bayesian generalized linear models. These methodologies should be seriously considered, especially since computational difficulties are dramatically reduced with the advancement of the MCMC methodology. This methodology allows us, for instance, to do away with empirical Bayes methods, correctly criticized by de Alba, and to carry full Bayesian analysis via hierarchical models.

Finally, Scollnik brought our attention to his recent paper on MCMC, which referred to the BUGS software as well as many additional recent research works in actuarial science incorporating Bayesian methods.