

Table of Contents

The CAS Institute Special Issue

4 **A Note from the Editor** *by Richard Gorvett*

6 **Contributors to this Issue**

13 **Minimum Bias, GLMs, and Credibility in the Context of Predictive Modeling**

by Chris Gross and Jon Evans

When predictive performance testing, rather than testing model assumptions, is used for validation, the need for detailed model specification is greatly reduced. Minimum bias models trade some degree of statistical independence in data points in exchange for statistically much more tame distributions underlying individual data points. A combination of multiplicative minimum bias and credibility methods for predictively modeling losses (pure premiums, claim counts, average severity, etc.) based on explanatory risk characteristics is defined. Advantages of this model include grounding in long-standing and conceptually lucid methods with minimal assumptions. An empirical case study is presented with comparisons between multiplicative minimum bias and a typical generalized linear model (GLM). Comparison is also made with methods of incorporating credibility into a GLM

39 **Embedded Predictive Analysis of Misrepresentation Risk in GLM Ratemaking Models** *by Michelle Xia, Lei Hua, and Gary Vadnais*

Misrepresentation is a type of insurance fraud that happens frequently in policy applications. Due to the unavailability of data, such frauds are usually expensive or difficult to detect. Based on the distributional structure of regular ratemaking data, we propose a generalized linear model (GLM) framework that allows for an embedded predictive analysis on the misrepresentation risk. In particular, we treat binary misrepresentation indicators as latent variables under GLM ratemaking models for rating factors that are subject to misrepresentation. Based on a latent logistic regression model on the prevalence of misrepresentation, the model identifies characteristics of policies that are subject to a high risk of misrepresentation. The method allows for multiple factors that are subject to misrepresentation, while accounting for other correctly measured risk factors. Based on the observed variables on the claim outcome and rating factors, we derive a mixture regression model structure that possesses identifiability. The identifiability ensures valid inference on the parameters of interest, including the rating relativities and the prevalence of misrepresentation. The usefulness of the method is demonstrated by simulation studies, as well as a case study using the Medical Expenditure Panel Survey data.

59 **Bayesian Predictive Modeling for Exponential-Pareto Composite Distribution** *by M. S. Aminzadeh and Min Deng*

Composite distributions have well-known applications in the insurance industry. In this paper, a composite exponential-Pareto distribution is considered, and the Bayes estimator under the squared error loss function is derived for the parameter θ , which is the boundary point for the supports of the two distributions. A predictive density is constructed under an inverse gamma prior distribution for the parameter θ , and the density is used to estimate the value at risk (VaR). Goodness of fit of the composite model is verified for a generated data set. The accuracy of the Bayes and VaR estimates is assessed via simulation studies. The “best” value for hyperparameters of the inverse gamma prior distribution are found via an upper bound on the variance of the prior distribution. Simulation studies indicate that when the “best” values of hyperparameters are used in the Bayes estimator, the estimator is consistently more accurate than maximum likelihood estimation.

69 **Machine Learning Methods to Perform Pricing Optimization. A Comparison with Standard GLMs** *by Giorgio Alfredo Spedicato, Christophe Dutang, and Leonardo Petrini*

As the level of competition increases, pricing optimization is gaining a central role in most mature insurance markets, forcing insurers to optimize their rating and consider customer behavior; the modeling scene for the latter is one currently dominated by frameworks based on generalized linear models (GLMs). In this paper, we explore the applicability of novel machine learning techniques, such as tree-boosted models, to optimize the proposed premium on prospective policyholders. Given their predictive gain over GLMs, we carefully analyze both the advantages and disadvantages induced by their use.

90 **On Prediction Of Future Insurance Claims When The Model Is Uncertain**

by Liang Hong, Todd Kuffner, and Ryan Martin

Predictive modeling is arguably one of the most important tasks actuaries face in their day-to-day work. In practice, actuaries may have a number of reasonable models to consider, all of which will provide different predictions. The most common strategy is first to use some kind of model selection tool to select a “best model” and then to use that model to make predictions. However, there is reason to be concerned about the use of the classical distribution theory to develop predictions because this theory ignores the selection effect. Since accuracy of predictions is crucial to the insurer’s pricing and solvency, care is needed to develop valid prediction methods. This paper investigates the effects of model selection on the validity of classical prediction tools and makes some recommendations for practitioners.

100 **Pricing Catastrophe Excess Of Loss Reinsurance Using Power Curves And The Generalized Logarithmic Mean** *by J. F. Walhin*

This paper advocates use of the generalized logarithmic mean as the midpoint of property catastrophe reinsurance layers when fitting rates on line with power curves. It demonstrates that the method is easy to implement and overcomes issues encountered when working with usual candidates for the midpoint, such as the arithmetic, geometric, or logarithmic mean. The paper shows how to deal with paid reinstatements in a simplified framework and also introduces a new midpoint that is consistent with a negative exponential fit of the rates on line.

A Note *from the Editor*



A Special Issue of *Variance* for an Evolving Profession

Welcome to the first “special issue” of *Variance*! In this issue appear six papers related to the topics of the first two programs being developed by iCAS, The CAS Institute: predictive analytics and catastrophe modeling.

It’s interesting to note that each of the papers here would be publishable in *Variance* even without the existence of iCAS. Indeed, the papers were all submitted to our journal for general consideration—they were not submitted as part of a special call for papers—and were reviewed and accepted on that basis. This is significant because it demonstrates the evolution of our casualty actuarial profession.

According to our website, “*Variance* is a peer-reviewed journal published by the Casualty Actuarial Society to disseminate work of interest to casualty actuaries worldwide. The focus of *Variance* is original practical and theoretical research in casualty actuarial science.” That these papers belong in *Variance* demonstrates the many ways in which actuarial science and, for example, data and predictive analytics have come to intersect. Some manifestations of this development:

- **Nature of the actuarial profession:** Actuaries have always been “data scientists” to some degree. As data has accumulated exponentially and become more accessible, and as computation speed and efficiency have exploded, the data analytics aspect of actuarial work has become more prevalent and important. Add to an actuary’s technical skills an understanding of the insurance and risk domain, and you have an ideal recipe for data science.
- **Actuarial exams:** The CAS’s newest professional exams, Modern Actuarial Statistics (MAS) I and II, were designed in reflection of the increasing importance, prevalence, and necessity of data and predictive analytics in actuaries’ lives.
- **iCAS:** The CAS Institute is a subsidiary of the CAS, and its first credential (CSPA) was created, in part, to support and acknowledge the critical role that data analytics and predictive modeling are now playing in the careers of many actuaries.

So it’s no accident that, among the general submissions to *Variance*, we easily found a half-dozen strong papers in areas relevant to the credentials being developed and offered by iCAS, and decided that, together, they would make a nice special issue.

More specifically, here's what you'll find in this issue regarding predictive modeling and analytics:

- **Gross and Evans** describe a combination of minimum bias and credibility methods for predictively modeling losses (pure premiums, claim counts, and/or average severity, etc.) based on explanatory risk characteristics, and provide an empirical case study for comparisons with GLM approaches.
- **Xia, Hua, and Vadnais** propose a GLM framework that allows for an embedded predictive analysis on misrepresentation risk. The usefulness of the method is demonstrated by simulation studies, as well as a case study using the Medical Expenditure Panel Survey data.
- **Aminzadeh and Deng** explore a composite exponential-Pareto distribution, and assess the accuracies of Bayes and other predictive estimators via simulation studies.
- **Spedicato, Dutang, and Petrini** explore the applicability of novel machine learning techniques to optimize the proposed premium on prospective policyholders, and analyze both the advantages and disadvantages of their use.
- **Hong, Kuffner, and Martin** undertake an investigation of the effects of model selection on the validity of classical prediction tools for insurance claims, and make some recommendations for practitioners.

In addition to these five papers related to predictive modeling, we present one paper related to catastrophe modeling, the subject of the second iCAS credential:

- **Walhin** advocates for the use of the generalized logarithmic mean as the midpoint of property catastrophe reinsurance layers when fitting rates on line with power curves, and addresses implementation and other issues.

Again, welcome to the first special issue of *Variance*! With the continuing evolution of our profession, I dare to predict that it won't be the last...

Rick Gorvett, editor in chief, *Variance*

Contributors

to this Issue



M. S. Aminzadeh

Dr. Aminzadeh is a 1985 Ph.D. in Statistics from Department of Statistics, Oregon State University. Before joining the Mathematics Department/Towson University and Johns Hopkins University in 1988, he taught in the Department of Statistics at Temple University (1985-1988). He is currently a full professor at Towson University and a graduate lecturer at Applied Physics lab (APL), Johns Hopkins University. He has taught a variety of applied and theoretical courses in mathematics, statistics, and stochastic processes. Several of his research papers were published in *Applied Statistics*, *Statistical Computation and Simulation*, *Computational Statistics*, *IEEE Transactions on Reliability*, and *Communications in Statistics*, and *Theory and Methods*.



Min Deng

Dr. Min Deng, FSA, received her Ph.D. from Pennsylvania State University in 1990. She has over 25 years of distinguished teaching experience, specializing in mathematics, statistics, and actuarial science. Her teaching methods have been widely recognized through multiple excellent teacher awards at both University of Wisconsin–Stevens Point and Maryville University. She has widely published in research journals. She is the author of the book *Paper: An Engineered Stochastic Structure* and many research papers. Her research interests include engineered stochastic structure, risk theory, and model construction. Currently she is a professor at Towson University and the director of ASRM program.



Christophe Dutang

Christophe Dutang is an assistant professor at University Paris Dauphine, part of PSL research university and is responsible for the actuarial program in Dauphine. He is a certified member of the French Institute of Actuaries and a member of the French Mathematical Society. He holds a Ph.D. from the University of Lyon. Christophe's research interests deal with stochastic reserving methods, parametric model for non-life insurance pricing, and the use of game theory in actuarial science. During the last years, he has given talks in 20 international conferences and published 13 papers in peer-reviewed journals. He is also reviewer for many scientific journals, an associate editor of the *Journal of Statistical Software*.

Editorial Board

Richard Gorvett
Editor in Chief

Sean P. McDermott
Ex Officio

Roger Bovard
Dale Edlefsen

Louise Francis
Richard W. Gorvett
John Major
Donald Mango
David L. Ruhm
Greg Taylor
Gary Venter

ASSISTANT EDITORS

Joel Atkins
Jean-Philippe Boucher
Frank H. Chang
Richard W. Gorvett

Daniel Heyer
Dmitry Papush
George Turner
Alexandros Zimbidis

Contributors

to this Issue



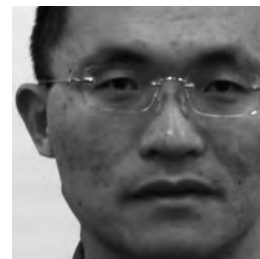
Jonathan Evans

Jonathan Evans, FCAS, FSA, FCA, CERA, MAAA, WCP is a property and casualty consulting actuary with over 20 years of experience. He is currently president of Convergent Actuarial Services, Inc. His writing includes 16 research papers and articles on actuarial and insurance subjects, and dozens of puzzles for the *Actuarial Review*.



Christopher Gross

Chris Gross, ACAS, MAAA is President and CEO of Gross Consulting in St. Paul, Minnesota. His primary areas of research are in predictive analytics, the use of claim-level development history to estimate insurance company reserves, methods to appropriately reflect differences in claim development when building insurance pricing models, and the construction of alternative case reserve estimates for use in actuarial reserve and pricing calculations. He is the chief architect of Gross Consulting's various software products. He has a B.S. in Economics and a Bachelor of Mathematics from the University of Minnesota.



Liang Hong

Liang Hong is a Fellow of the Society of Actuaries and an associate professor in the Department of Mathematics at Robert Morris University. He received his PhD from Department of Mathematics at Purdue University in 2009. Liang's research interests include Bayesian and Bayesian-like methods, credibility theory, optimal insurance level, and predictive modeling.

COPY EDITORS

Nathan J. Babcock
Hsiu-Mei Chang
Andrew Samuel Golfin Jr.
Mark Komiskey
William E. Vogan

CAS STAFF

Elizabeth A. Smith
Manager of Publications
Donna Royston
Publications Production Coordinator
Sonja Uyenco
Desktop Publisher



Contributors

to this Issue



Lei Hua

Lei (Larry) Hua is an Associate Professor of Actuarial Science at Northern Illinois University. He got a PhD degree in Statistics from the University of British Columbia. He is an Associate of the Society of Actuaries. His research work focuses on multivariate dependence modeling for non-Gaussian phenomena and innovative applications for financial and insurance industries. His current interest is in cutting-edge machine learning techniques, including deep learning for dealing with high frequency financial trading data. In his spare time, Larry enjoys spending time with his kids and running marathons.



Todd Kuffner

Todd Kuffner is an Assistant Professor in the Department of Mathematics and Statistics at Washington University in St. Louis, and his Ph.D. is from the Department of Mathematics at Imperial College London in 2011. Todd's research interests include higher-order asymptotics, post-selection inference, the bootstrap, and Bayes-frequentist reconciliation.



Ryan Martin

Ryan Martin is an Associate Professor in the Department of Statistics at North Carolina State University, and his PhD is from the Department of Statistics at Purdue University in 2009. Ryan's research interests include asymptotics, Bayes and empirical Bayes analysis, foundations of statistics, high- and infinite-dimensional problems, and mixture models. He is co-author of the monograph *Inferential Models*, published by Chapman-Hall/CRC Press in 2016.

Variance is published twice yearly by the Casualty Actuarial Society. Telephone: (703) 276-3100; Fax: (703) 276-3108; E-mail: office@casact.org. Presorted Bound Printed Matter postage is paid at Baltimore, Maryland. Publications Mail Agreement No. 40035891. Return Undeliverable Canadian Addresses to PO Box 503, RPO West Beaver Creek, Richmond Hill, ON L4B 4R6.

Postmaster: Send address changes to: *Variance*, 4350 North Fairfax Drive, Suite 250, Arlington, Virginia 22203.

For permission to reprint material from *Variance*, please write to the editor in chief. Letters to the editor can be sent to staffeditor@variancejournal.org or the CAS Office. The Casualty Actuarial Society is not responsible for statements or opinions expressed in the articles, discussions, or letters printed in *Variance*.

© 2018 Casualty Actuarial Society.

Mailing Address

Variance

4350 North Fairfax Drive
Suite 250

Arlington, Virginia 22203

staffeditor@variancejournal.org

For information on
submitting papers visit

www.variancejournal.com.

Contributors

to this Issue



Leonardo Petrini

Leonardo is currently COO and data scientist at Virality. Recently, he has applied his statistics and machine learning knowledge to several fields such as insurance, finance, and social media, with a focus on computer vision problems. He's engaged in solving problems in pricing optimization in insurance and deep learning for computer vision, as well as finding highly scalable natural language solutions.



Giorgio A. Spedicato

Giorgio A. Spedicato is a Senior Data Scientist at the Unipol Group, Italy. Before joining the Unipol Group, he worked in P&C pricing, reserving and economic capital roles for the Italian branches of the Axa and Aviva groups and as statistical consultant. He holds a PhD in actuarial science, the SOA and CAS fellowships, and the CSPA credential. He is author of some research papers in peer-reviewed journals.



Gary Vadnais

Gary Vadnais, FCAS, is the capital manager for the Canadian-based Intact Financial Corporation. He taught secondary mathematics for three years prior to joining Intact in 2003. Since then, he has been a contributor to the evolution and modernization of internal pricing models and managed national reserving, and now focuses on capital optimization. He currently sits on national industry committees, consulting on changes to regulatory capital policies for Canada.



Contributors

to this Issue



Jean-François Walhin

Jean-François Walhin is CEO of Aon Reinsurance Solutions in Brussels, in charge of Belgium and Luxembourg. Jean-François joined Aon in 2009 and was previously Chief Underwriting Officer of Fortis Re as well as head of non-life and reinsurance of Fortis Insurance. Prior to that Jean-François worked 7 years with QBE Re and 3 years with MMA in Brussels. Jean-François has over 20 years of reinsurance experience as a cedant, a reinsurer, a captive manager, and a broker. Jean-François is a civil engineer, a fellow of the IA|BE, holds a PhD in Statistics (UCL) and has a MBA from the Vlerick School. He is also a visiting professor at UCL (Louvain-la-Neuve) and ULB (Brussels).



Michelle Xia

Michelle Xia is the Director of Statistical Consulting Services and an Assistant Professor in Statistics at Northern Illinois University (NIU). Michelle obtained her PhD in statistics at the University of British Columbia, her master's in actuarial science from the University of Calgary, and dual bachelor's degrees in insurance and mathematical statistics and probability from Nankai University. Besides research, teaching and consulting at NIU, Michelle has over seven years of professional experiences as an actuary, a predictive modeler and a statistician in insurance and medical areas.