actuarial REVIEW

VOL 44 / NO 3 / MAY-JUNE 2017
PUBLISHED BY THE CASUALTY ACTUARIAL SOCIETY

HOW PREDICTIVE IS THE PAST WHEN FUTURE RISK IS UNFATHOMABLY DYNAMIC?

2017 CAS PREDICTIVE ANALYTICS MARKETPLACE

ACTUARIAL CAREERS, INC.®

2016

Salary Survey

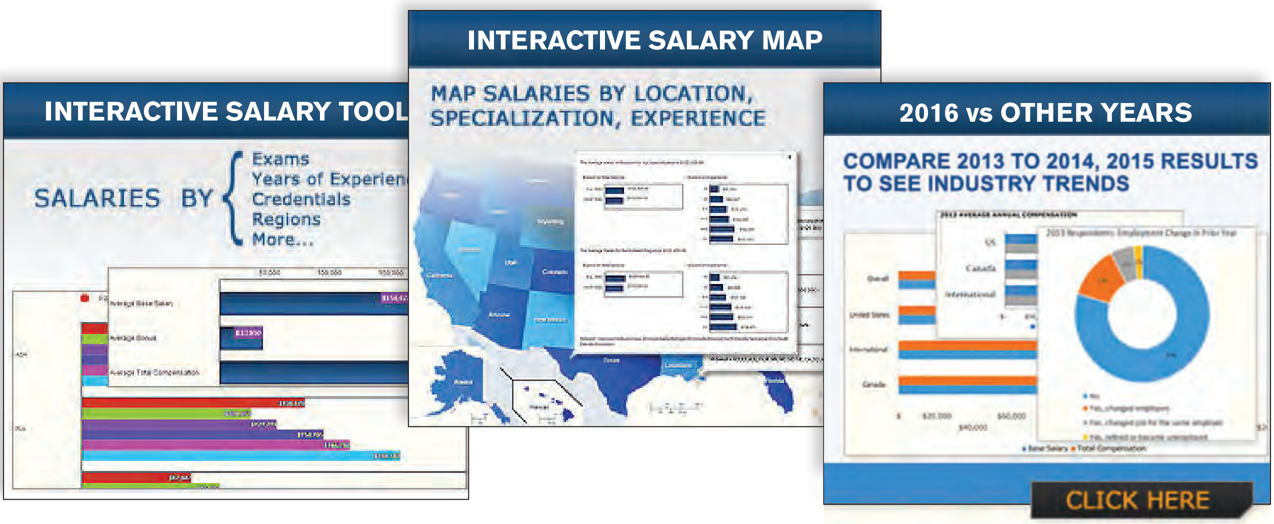
Our 2016 Salary Survey results are here!

Are you curious about actuarial salaries? NOW is the time to go online to www.actuarialcareers.com/salary-survey/ to access our 2016 salary survey results. You can run queries on the results and see where you fall on the industry salary scale.

Our online query tools allow you to select and display information that is pertinent to earnings in an array of combinations including: Specialization, Experience, Education and Location.

This year our results represent responses to questionnaires we sent to more than 40,000 actuaries, others who volunteered to participate, and from information we gather from candidates and the companies we recruit for.

There are a few samples below, but you must go to our website http://www.actuarialcareers.com/ and click on the Salary Survey tab to find the 2016 results. You can also see and query past year's results too!



ACTUARIAL CAREERS, INC.®
11 Martine Avenue, 9th Floor / White Plains, NY 10606 / Tel: 914-285-5100 / Toll Free: 800-766-0070 / Fax: 914-285-9375
www.actuarialcareers.com / E-mail: jobs@actuarialcareers.com

actuarialREVIEW

May/June 2017

departments

4 **EDITOR'S NOTE**
- Building a Legacy

6 **PRESIDENT'S MESSAGE**
- Research in a World of Change

8 **MEMBER NEWS**
- Calendar of Events
- Comings and Goings
- In Memoriam
- CAS Staff Spotlight
- CAS Takes Part in Second Annual Insurance Careers Month
- Certified Specialists in Predictive Analytics Recognized in March 2017
- Scenes from the CAS RPM Seminar and Workshops
- CAS Student Ambassador Program and CLRS Livestream Event Honored With Industry Awards

44 **PROFESSIONAL INSIGHT**
- Darwinian Theory Meets Insurance Analytics
- Advanced Analytics: Getting Under Way
- You've Built the Model. Have You Done It Right?

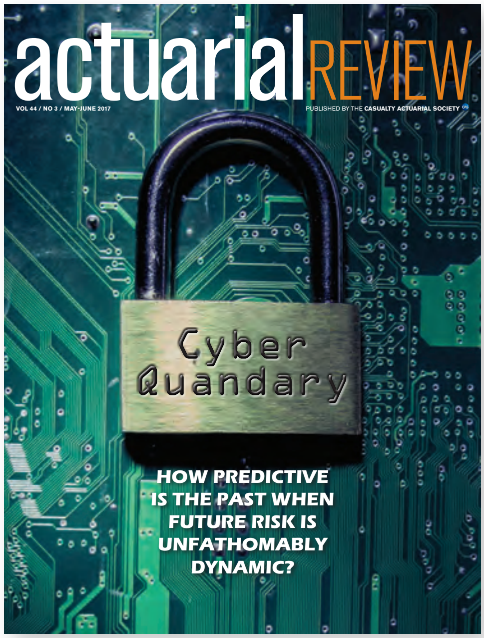
52 **ACTUARIAL EXPERTISE**
- Explorations

54 **VIEWPOINT**
- In My Opinion

56 **SOLVE THIS**
- It's a Puzzlement

FSC LOGO

on the cover



BY ANNMARIE GEDDES BARIBEAU 37

2017 CAS Predictive Analytics Marketplace

A Supplement to *Actuarial Review* 21

Actuarial Review (ISSN 10465081) is published bimonthly by the Casualty Actuarial Society, 4350 Fairfax Drive, Suite 250, Arlington, VA 22203. Telephone: (703) 276-3100; Fax: (703) 276-3108; Email: ar@casact.org. Presorted standard postage is paid in Lutherville, MD. Publications Mail Agreement No. 40035891. Return Undeliverable Canadian Addresses to PO Box 503, RPO West Beaver Creek, Richmond Hill, ON L4B 4R6.

The amount of dues applied toward each subscription of *Actuarial Review* is $10. Subscriptions to nonmembers are $10 per year. Postmaster: Send address changes to *Actuarial Review*, 4350 North Fairfax Drive, Suite 250, Arlington, Virginia 22203.

editor's NOTE By ELIZABETH SMITH, *AR* MANAGING EDITOR

Building a Legacy

Grover Edie's "In My Opinion" column got me thinking about legacies.

Many of us have children and some of us have built businesses or created art. These and others are all fine accomplishments, but they are not out of the ordinary.

One of us is not likely to write the great American novel — or the Canadian or Chinese or Danish one for that matter. For the most part, we're pretty good people — some would even say that we're impressive — but only a select few will have monuments built in their honor. We may not all leave records of our time here on Earth.

The thing is, a legacy is sometimes not a "thing" at all – at least not something you can touch, but rather something that is felt.

A legacy can be something quite simple. Take, for example, a recipe passed on from parent to child. The recipe is not just a list of ingredients and a series of steps: it's the time spent together, the aroma of the ingredients cooking and the taste of the end result. All of these "things" can combine to create a happy experience and, later, a warm memory.

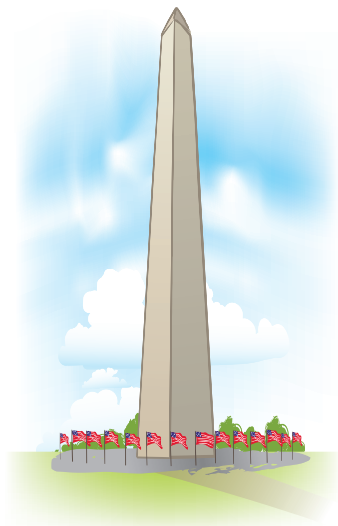
And so, a legacy is really quite simple: It's time.

It's the time spent with loved ones, teaching them a skill or taking them along with you on a business trip. It's the time spent with a young person, offering them career advice and guidance.

Who needs monuments when you can make good memories?

So I ask you: Will you take the time to build *your* legacy?

P.S. Following up on Grover Edie's prior "In My Opinion" column (*AR*, March-April 2017), the U.S. Patent and Trademark Office approved both of his trademark renewals!



Actuarial Review always welcomes story ideas from our readers. Please specify which department you intend for your item: Member News, Solve This, Professional Insight, Actuarial Expertise, etc.

SEND YOUR COMMENTS AND SUGGESTIONS TO:
Actuarial Review
Casualty Actuarial Society
4350 North Fairfax Drive, Suite 250
Arlington, Virginia 22203 USA
Or email us at AR@casact.org

Follow the CAS

president'sMESSAGE By NANCY BRAITHWAITE

Research in a World of Change

Continuing the theme of change from my last column, it's a good time to talk about how the CAS's research efforts fit in to the overall picture. Remember, we are dedicated to the advancement of casualty actuarial science. This advancement involves not only education through Fellowship and beyond, but also encouraging and targeting research to improve our success in providing useful insights to our business partners.

As the world around us changes, the circumstances around the financial outcomes we are trying to estimate are changing as well. Think of cyberrisk exposures. Are our traditional methods appropriate as we try to estimate the financial impacts of events that we are just beginning to imagine? As our cars and homes get "smarter" are we creating potential exposure to different kinds of losses? What are these losses and what will they cost?

As data proliferates around us and methods for analyzing this data change, are we thinking about and testing new methods for estimating liabilities? Are our reserving and pricing methods becoming more sophisticated? Is our risk selection improving?

If we are truly dedicated to continuous improvement, we must be thinking about ways to encourage further research and publishing cutting-edge knowledge. Over the last few years, we have been working on bringing our research committees forward to become more supportive of innovation. We are engaging in research working parties (similar to task forces) that are focused on results. We are developing communities of interest, so members can share their ideas and build off each other. We are focusing on engaging with younger members in the research areas to keep current and keep the flow of new ideas coming.

We've identified five areas of research focus for the near term. (Of course, this doesn't mean that other areas of research will be ignored.) These are the areas we believe are most affected by the significant changes we are seeing around us:

- Predictive modeling and data analytics.
- Modeling, more broadly.
- Reserving methods, processes and validation.
- Economic scenarios and stress-testing.
- Cyberrisk.

Think of cyberrisk exposures. Are our traditional methods appropriate as we try to estimate the financial impacts of events that we are just beginning to imagine?

Predictive analytics and modeling take a prominent place in this list. Obviously, there is a lot of buzz in the marketplace. We need to remember that, in many instances, we are working in a regulated environment. We need some serious research and discussion around how to make the best use of our new tools while remaining within the letter and the spirit of the regulation around us. We are exploring the creation of a working party on the ethical, legal and regulatory considerations in using predictive models.

To continue to encourage innovative thought in these and other areas, we are sponsoring funded research and collaborating with other actuarial organizations around the world on scientific investigations. We continue to form working parties and conduct call paper programs, and we encourage our members to get involved.

We are the foremost actuarial organization in the world in property-casualty practice, and if we want to remain the leader, we need to continue to lead. For that, we need to continue to be thoughtful and innovative in developing practical approaches and solutions to new and old problems. We need to be constantly moving forward in our world of change.

CALENDAR OF EVENTS

May 21 - 24, 2017
Spring Meeting
Sheraton Centre Toronto Hotel
Toronto, ON

June 5 - 6, 2017
Seminar on Reinsurance
Fairmont Washington, DC
Washington, DC

September 10 - 12, 2017
Casualty Loss Reserve Seminar (CLRS) & Workshops
Loews Philadelphia Hotel
Philadelphia, PA

November 5 - 8, 2017
Annual Meeting
Fairmont Austin
Austin, TX

March 19 - 21, 2018
Ratemaking and Product Management (RPM) Seminar & Workshops
Fairmont Chicago, Millennial Park
Chicago, IL

May 13 - 16, 2018
Spring Meeting
Boston Marriott Copley Place
Boston, MA

June 4 - 5, 2018
Seminar on Reinsurance
New York Marriott at the Brooklyn Bridge
New York, NY

COMINGS AND GOINGS

Angela Burgess, FCAS, MAAA, was recently named as senior vice president and chief actuary of Assurant, Inc., a Fortune 500 global provider of specialty insurance products and risk management solutions in the housing and lifestyle markets. Burgess, based in Miami, Florida, oversees a team of 100 actuaries worldwide and serves as the Miami CAS exam site coordinator.

Paul G. O'Connell, FCAS, MAAA, has been promoted to senior vice president and chief actuary at Chubb Group. O'Connell will oversee all actuarial functions, including reserving, pricing and capital performance measurement. Prior to ACE's acquisition of Chubb in January 2016, O'Connell was chief actuary, global property and casualty for ACE, a position he held since 2010. Prior to joining ACE in 2002, he was a principal at PricewaterhouseCoopers.

QBE North America has appointed **John A. Beckman, FCAS**, as its new chief underwriting officer. Beckman will be responsible for leading the company's efforts to enhance existing underwriting practices and processes, including product development, technical leadership and risk appetite. Beckman most recently served as senior vice president and chief transformation officer at CNA Insurance. Beckman was the chief underwriting officer of CNA's commercial insurance business from 2011 to 2015.

Uri Korn, FCAS, MAAA, is the 2017 recipient of the CAS Ratemaking Prize for his paper "An Alternative Approach to Credibility for Large Account and Excess of Loss Treaty Pricing." Korn is the industry analytics leader for AIG Client Risk Solutions. His work and research experience includes practical applications of credibility, trend estimation, increased limit factors, non-aggregated loss development methods and Bayesian models. Korn's paper is posted in the 2017 Spring *E-Forum*.

EMAIL "COMINGS AND GOINGS" ITEMS TO AR@CASACT.ORG.

ACTUARIAL REVIEW *LETTERS POLICIES*

Letters to the editor may be sent to ar@casact.org or the CAS Office address. Include a telephone number with all letters. Actuarial Review reserves the right to edit all letters for length and clarity and cannot assure the publication of any letter. Please limit letters to 250 words. Under special circumstances, writers may request anonymity, but no letter will be printed if the author's identity is unknown to the editors. Announcement of events will not be printed.

IN MEMORIAM

Daniel A. Crifo (ACAS 1977)
1947-2017

Daniel J. Flaherty (FCAS 1966)
1941-2017

John J. Kollar (FCAS 1975)
1947-2017

John W. "Bill" Wieder Jr. (FCAS 1947)
1918-2017

SPRING MEETING

May 21-24, 2017
Sheraton Centre Toronto Hotel
Toronto, Canada

IMAGINE: CONFIDENCE IN THE NUMBERS.


IT TAKES VISION
Introducing Arius,® the state of the art in reserving solutions.
Designed from the ground up by Milliman, Arius delivers proven innovations like deterministic analysis combined with advanced variability models, all in a customizable work environment. The bottom line? A better understanding of your numbers, and smarter business decisions. So say goodbye to patchwork spreadsheets, and hello to Arius. Milliman.com/Arius
ARIUS®
BUILT BY Milliman

CAS STAFF SPOTLIGHT

Meet Sonja Uyenco, Desktop Publisher

Welcome to the CAS Staff Spotlight, a column featuring members of the CAS staff. For this spotlight, we are proud to introduce you to Sonja Uyenco.

- **What do you do at the CAS?**
  I lay out the very publication you are reading right now! My work also includes *Future Fellows* and the postcards, registration brochures and onsite programs for CAS meetings and seminars. Pretty much if it is printed, I designed it.
- **What do you enjoy most about your job?**
  I enjoy coming up with designs and the challenge of making a layout fit together like a Tetris puzzle, especially when a small change completely rearranges a whole layout.
- **What's your hometown?**
  Fort Washington, Maryland.
- **Where'd you go to college and what's your degree?**
  B.A. in Studio Art with Graphic Design concentration at University of Maryland — Go Terps! My favorite animal happens to be the turtle, but that didn't influence my decision to go to UM.
- **What was your first job out of college?**
  Graphic design specialist at Corporate Executive Board.
- **Describe yourself in three words:**
  Introverted adrenaline junkie.
- **What's your favorite weekend activity?**
  During the winter it's snowboarding, but for the rest of the year it depends. I actually love staying home and having nothing planned on weekends because that is when I recharge from doing most of my other favorite activities, like weightlifting, capoeira and archery, during the week.

Sonja Uyenco

- **What's your favorite travel destination?**
  Hawaii. My grandmother lives there, so I visit often. It would be perfect if I knew how to surf, but I panic easily in open water, so I stick with snorkeling beside the sea turtles.
- **Name one interesting or fun fact about you:**
  Despite all my tattoos, I'm scared of IV needles.

CAS Takes Part in Second Annual Insurance Careers Month

BY TAMAR GERTNER, CAS UNIVERSITY ENGAGEMENT MANAGER

This past February, the CAS participated in the second annual Insurance Careers Month (ICM) campaign, alongside over 850 other insurance organizations that recognize the importance of attracting the next generation of leaders to the insurance industry. The campaign's theme is that insurance represents the "Career Trifecta" — stable, rewarding and limitless.

ICM was initiated in 2016 by Hamilton Insurance Group and its chairman and CEO Brian Duperreault, ACAS, MAAA, along with a handful of other organizations. InsuranceCareersTrifecta.org is the campaign's website and provides participating organizations with access, resources and tools to support their involvement.

Leading up to the kick-off to ICM 2017, a CEO Town Hall webinar was held to educate insurance professionals and organizations about the campaign. More than 1,000 participants from the United States, Canada, Bermuda and the United Kingdom heard insurance company CEOs discuss the industry's talent gap. Duperreault explained how critical it is to change the current image of the insurance industry to attract the next generation of insurance industry professionals. "Insurance should be catnip to millennials looking for a purpose-driven career. What other industry provides protection and security, or supports growth or recovery from disaster?"

Duperreault said that it is the job of the industry to explain what it does and why it matters. "We will be facing gaps in management and new talent that make it difficult, if not impossible, to meet the needs of the digital world we are living in. With the sharing economy, technological advances and associated emerging risks, this is the brave new world in insurance." He implored the webinar participants to help spread the word. "You've chosen a career in insurance, so yours is the best voice for other millennials to hear," said Duperreault.

Posted on CAS and Student Central social media platforms throughout the month of February, the Millennial Profile series featured millennial CAS members and candidates offering their insights and advice to students on careers in the insurance industry.

(Visit InsuranceCareersTrifecta.org to view the webinar and hear the full set of comments.)

Campaign supporters answered the call. During ICM in February, the number of companies participating grew from 600 in 2016 to 850 in 2017. Together, they created a strong online presence, using social media hashtags #CareerTrifecta and #TalentTuesday almost 3,000 times on Twitter, amounting to 10.9 million impressions throughout the month.

For the second year in a row, the CAS ran a Millennial Profile series in February in which millennial CAS members and candidates provided insights and advice for students about the insurance industry. Featured millennials shared why they became actuaries, why they'd recommend a career in the insurance industry, and what students can do now to prepare for a career in the industry. The profiles were circulated on CAS and Student Central social media platforms and can be viewed at CASstudentcentral.org.

In addition, the CAS supported the movement by coordinating and encouraging activity among a coalition of insurance industry trade and professional associations, led by CAS Chief Communications Officer Mike Boa.

Other insurance organizations participated in a variety of ways. Gamma Iota Sigma, the premier talent pipeline to the insurance industry and sole international business fraternity for students of insurance, risk management and actuarial science, held two live tweet-chats as part of their month of many supporting activities. Swiss Re spread the word with a blog post and a video that encourages millennials to rethink the insurance industry. Hamilton Re launched a contest for Bermuda high school students, giving them a chance to learn what it's like to work in the reinsurance industry. Millennials across companies were highlighted in articles, including CNA's Eric Blancke, FCAS, an actuarial consultant featured in a LinkedIn article about the diversity of his work.

Insurance Careers Month and the broader year-round movement align well with the CAS's extensive ongoing university engagement efforts promoting the P&C career to millennials.

A dedicated network of more than 300 CAS members worldwide volunteer to bring awareness to university students about P&C careers through the CAS University Liaison Program. The CAS also offers a free student membership program, CAS Student Central, providing students access to actuarial exam prep materials, P&C internship listings, scholarship information and webinar and networking event invitations.

Visit InsuranceCareersTrifecta.org for more information on how your company can get involved in the movement.

Insurance Careers Month 2017 was led by Hamilton Insurance Group, Lloyd's of London, Marsh & McLennan, The Institutes, MyPath, Valen Analytics, The Jacobson Group, InVEST and PCI.

CERTIFIED SPECIALISTS IN PREDICTIVE ANALYTICS RECOGNIZED IN MARCH 2017

Seated, left to right: *Louise Francis, Susan Poole, Guangjin (Jim) Xiao, Stephen Stone,* **iCAS Leadership Advisory Council Chair Robert Miccolis**, *Ravi Kumar, Todd Lehmann and Cheng-Sheng Peter Wu.*
Standing, left to right: CAS President-Elect Brian Brown, *Christopher Monsour, William Frierson, Jeffrey Kinsey, Hernan Medina, Trent Goughnour, Gregory Hayward, Andrew Sutcliffe. Photo credit: Crown City Photography.*

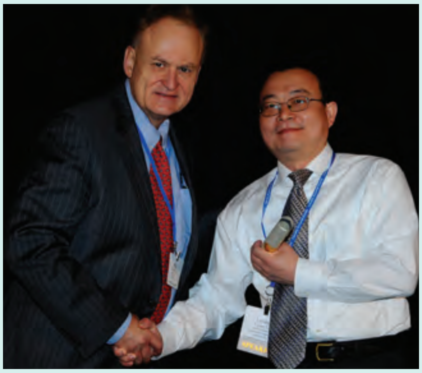
Not pictured are Avraham Adler, Joel Atkins, Shane Barnes, Andrew Brown, Richard Crabb, Denise Christophel, Linhui Dong, Luyang Fu, James Guszcza, Ronald Lettofsky, Weiting Lu, Zachary Martin, Stephen Mildenhall, Roosevelt Mosley, Ernesto Schirmacher, Rebecca Vessenes, Jonathan Zabek, Ya Zhang.

SEMINAR ON REINSURANCE

JUNE 5-6, 2017
THE FAIRMONT
WASHINGTON, DC

CAS

Scenes from the CAS RPM Seminar and Workshops



Luyang Fu, CSPA, FCAS (right), receives his CSPA certification from CAS President-Elect Brian Brown.

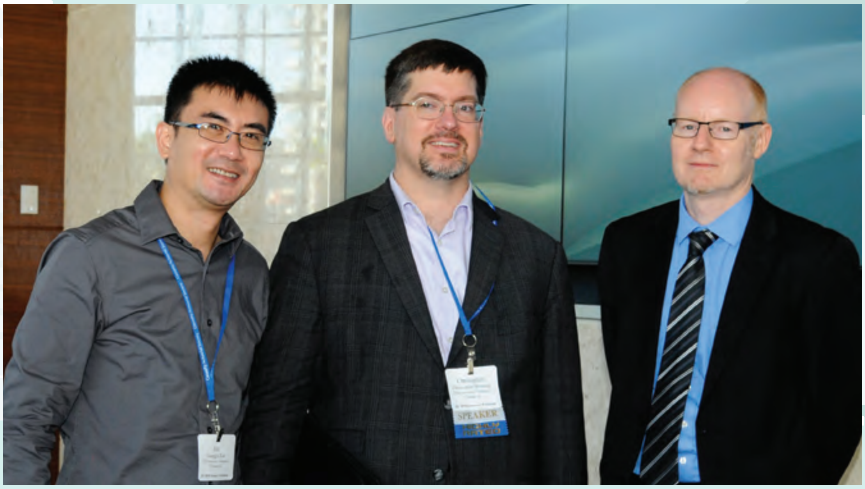
Author James P. Lynch, FCAS, (left) and Gary C. Wang, FCAS.

Chris Cooksey, FCAS (left) presents the 2017 CAS Ratemaking Award to Uri Korn, FCAS.

Sandra Callanan, FCAS, at the iCAS Predictive Analytics Community of Practice event on March 27.

Trevor Soupir, FCAS, listens during the concurrent session "Genetic Algorithms with Applications in Insurance," held on March 29.

RPM attendees at the concurrent session "How to Pick a Better Model" on March 28.



Adding another designation to their CVs are (left to right) Guangjin "Jim" Xiao, CSPA, FCAS, MAAA; Christopher Monsour, CSPA, FCAS, MAAA; and Andrew Sutcliffe, CSPA.

CAS Student Ambassador Program and CLRS Livestream Event Honored With Industry Awards

BY KATE NISWANDER, CAS MARKETING AND COMMUNICATIONS MANAGER

The CAS was recognized for outstanding membership and e-learning programs by the 2016 *Association TRENDS*' All Media Contest during the annual Salute to Association Excellence celebration in Washington, D.C., last March.

The annual competition acknowledges the most creative and effective communications for associations.

The 2016 Casualty Loss Reserve Seminar (CLRS) Livestream event received a gold award in the category of eLearning and Live Training. The CAS Student Central Ambassador program received a silver in the Membership Promotion category.

The winning entries were among more than 400 entries submitted in the 2016 competition.

The 2016 Casualty Loss Reserve Seminar (CLRS) Livestream event offered participants the opportunity to attend and interact remotely in six seminar sessions.

In 2015 the CAS began livestreaming select sessions from its conferences and seminars as a convenient way for members to affordably and conveniently earn continuing education credit.

Offering "hybrid" events (with both a virtual and face-to-face component) has enabled more members to participate in CAS professional education programs. The sessions live-streamed at the 2016 CLRS event ranged from hot topics — such as climate change and its effect on risk management — to more traditional actuarial topics — such as analytics and machine learning. The event also included an evening session aimed at the CAS's fast-growing international audience in Asia.

The CAS Student Central Ambassador program, piloted in 2015-16 and now in its second year, was launched in response to the continued growth of CAS Student Central, the free membership program for university students pursuing a career as an actuary. In order to increase the reach of CAS Student Central and continue building on-campus engagement, the CAS piloted the Ambassador Program at 12 colleges and universities. Twelve exceptional actuarial students served as CAS Ambassadors, with the goals of increasing the awareness of CAS Student Central among their classmates and strengthening CAS ties to the university. Ambassadors hosted on-campus events featuring CAS member speakers, provided information to fellow students, and spread the word about the actuarial profession as a career choice.

"It is an honor to once again be recognized by *Association TRENDS* for our member services and marketing and communications programs," said Mike Boa, CAS chief communications officer. "Our professional education department has developed a convenient way for members to participate in learning opportunities they would not otherwise be able to attend. Our university outreach efforts continue to expand with the addition of our student ambassadors; we embrace the opportunity to have these bright and talented individuals as new additions to the strong volunteer culture of the CAS."

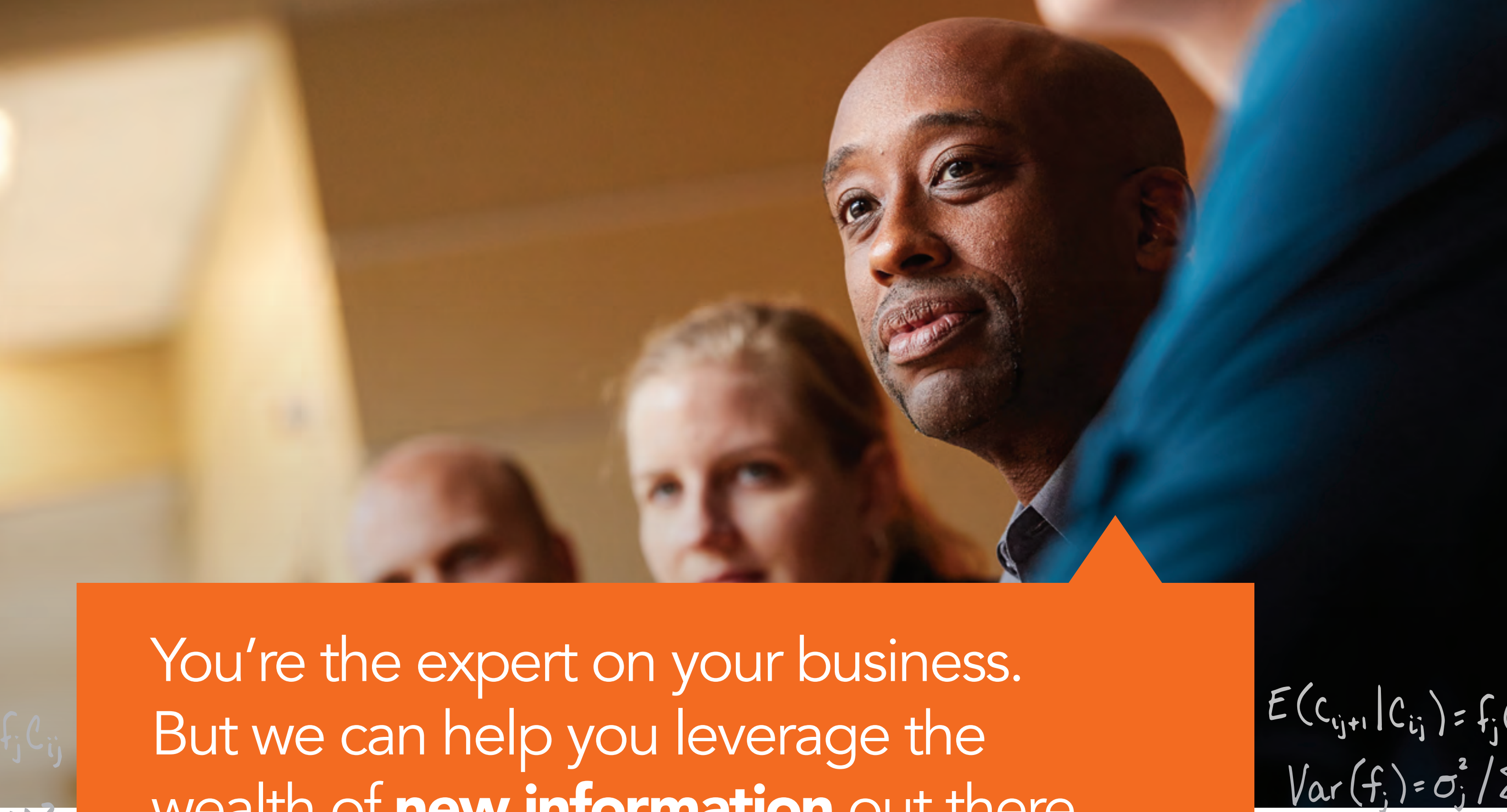
Association TRENDS is the national newspaper for association executives and suppliers, spotlighting the latest news, information and trends in association management for the professional staff of international, national, state, regional and local voluntary organizations.

For a complete list of 2016 All Media Contest winners, visit the competition website.

2017 CAS PREDICTIVE ANALYTICS MARKETPLACE

Expertise. Insight. Solutions.®

A SUPPLEMENT TO *ACTUARIAL REVIEW*



2017 CAS

PREDICTIVE ANALYTICS MARKETPLACE

Expertise. Insight. Solutions.®

TABLE OF CONTENTS

4 **Do Your Data Analytics Team Members Speak the Same Language?**
BY NANCY BRAITHWAITE, FCAS, MAAA, CPCU

6 **The CAS Institute Grants Its First Certified Specialist in Predictive Analytics (CSPA) Credentials**
BY KATE NISWANDER

8 **How Can Insurers Find Real Value in Their Predictive Models?**
BY CLAUDINE MODLIN, FCAS, MAAA

11 **Predictive Analytics: What's Next?**
BY ROOSEVELT C. MOSLEY JR., FCAS, MAAA, CSPA

13 **Added Values: Breathe Life into P&C Projections**
BY STEPHEN URBROCK

14 **Predictive Analytics Providers Directory**

Do Your Data Analytics Team Members Speak the Same Language?

BY NANCY BRAITHWAITE, FCAS, MAAA, CPCU

Each member of the data analytics team brings professional strengths but may not define terms the same way as another team member.

The world is constantly changing, and as an actuary, I probably view these changes differently than most people. In my world, all of the advancements in new and innovative technology that have made our lives more convenient also present more complex risks.

From an insurance perspective, new and more prevalent technologies like mobile payments and drones require more complex risk-management tools. Previous methods for quantifying and managing risk — such as using past data to price insurance products — may no longer be sufficient. At the same time, the digital revolution, led by smartphones and wearable devices, is giving us more data than ever before. Insurers need to embrace and mine the increasing volume of data, finding new techniques to evaluate and produce insights.

The good news is that a lot of new data is readily available — the not-so-good news is that insurers and their analytical teams may not know what to do with this data.

Everything from the sheer volume of data to the nature of how it is stored and processed can make it hard to sift through and find information that will be useful. Due to this influx of data, the industry has seen the partnering of actuarial work with data science to perform predictive modeling. Despite the arrival of new techniques, however, insurance remains a highly specialized and highly regulated industry — those handling the data in insurance companies need to fully understand the business context in which it lives. One variable of data may represent something that is not legal or socially acceptable to actually use in practice, or, data may say something that makes no sense at all — for example, that women with red hair have more auto accidents (when anyone can dye their hair). So those working on data analytics teams need to have a strong sense of causality when evaluating data, knowing how it plays into the larger business context of the problem they're trying to solve.

[T]hose working on data analytics teams need to have a strong sense of causality when evaluating data, knowing how it plays into the larger business context of the problem they're trying to solve.

Bridging the Communications Gap

There is no prescribed composition for an effective data analytics team — it can have a mix of data scientists, actuaries, statisticians, and others. Each professional brings something to the discussion, and increasingly the "team" approach to analytics results in success. However, those same professionals need to understand each other's perspectives — they need to be able to speak the same language in order to communicate and collaborate. Ideally, members of the team will have a certified set of predictive analytics skills, which can help set a standard and bridge the communication gap that exists.

For employers, this lack of common "language" in the predictive analytics environment can also affect their recruitment. Position titles such as "data scientist" or "modeler" do not have a consistent description or industry standard. Last year, when the Casualty Actuarial Society (CAS) conducted market research with insurance company executives on the subject, employers cited recruiting/hiring as one of their greatest challenges in predictive analytics. In fact, 76

Data
البيانات
شمار و اعداد کے
ข้อมูล
Datos
Dados
데이터
जानकारी
Daten
Les données
Տվյալներ
Podaci
ʻikepili
Gegevens
Δεδομένα
Gögn

percent of those surveyed noted that a certification would be beneficial to employers seeking to hire specialists in predictive modeling.

Becoming a Certified Predictive Analytics Specialist

This is one of the many reasons that The CAS Institute, a subsidiary of the CAS, recently launched its Certified Specialist in Predictive Analytics (CSPA) credential. The credential, created for data professionals with several to many years' experience, requires that candidates demonstrate evidence of applied knowledge in predictive analytics by passing a series of four assessments. The program draws from the history and strength of the CAS, whose high-quality educational standards and credentialing programs for actuaries have been recognized globally for over 100 years.

The curriculum of the CSPA credential is overseen by an expert panel comprising industry specialists working in predictive analytics. The four required assessments cover:

- The fundamentals of property and casualty insurance;
- How data works, including the forms it can take;
- How to present and work with data, including building models; and
- How to apply these skills to a real-life scenario.

The final assessment also asks candidates to complete data analysis and a report based on an assigned scenario. The candidate is required to integrate and apply all knowledge from the previous three assessments in order to achieve success.

Final projects will vary so as to reflect real-world-type predictive analytics scenarios. For example, one project might have candidates working to improve claims department operations, such as identifying potential high-severity claims, or controlling claims department costs. A marketing-focused project could ask candidates to improve sales through methods such as matching product offerings to customer type, or targeting new or optimal customer segments. CSPA candidates may also use their predictive analytics skills in scenarios involving underwriting, pricing, or even operations. This "case study" project helps round out the CSPA curriculum by testing the candidates' ability to use their predictive analytics skills in the workplace.

Employers are enthusiastic to see a program that can provide professional education and certification to members of their team who have previously been without these types of dedicated resources. Employers now have a reference point when they decide to add predictive analytics professionals to their staff.

A New Professional Community

CSPA credential holders are also required to complete an ethics course and adhere to a standard of professionalism and code of conduct, something not previously required of those in analytics roles.

After traveling all over the U.S. sharing information about our new CSPA credential with employers, we can say that the response has been overwhelmingly positive. Employers are enthusiastic to see a program that can provide professional education and certification to members of their team who have previously been without these types of dedicated resources. Employers now have a reference point when they decide to add predictive analytics professionals to their staff. The CAS Institute also provides its members with a professional community, where those working in this specialized field can connect.

Ultimately the expansion of predictive analytics within the insurance industry has opened doors for new opportunities to improve business performance. In order to maintain momentum and keep up with changes, predictive analytics teams need to make sure they are well-equipped and collaborating effectively to adapt to new technologies and new data. It's only through the improvement and standardization of analytical skills, coupled with the willingness to learn, that we will remain ready to respond to the technological (and societal) changes that still await us.

Nancy Braithwaite, FCAS, MAAA, CPCU, is a second vice president and actuary in the Excess Casualty Department at Travelers Insurance Co. She currently serves as president of the Casualty Actuarial Society (CAS). Opinions are the author's own.

The CAS Institute Grants Its First Certified Specialist in Predictive Analytics (CSPA) Credentials

BY KATE NISWANDER

The CAS Institute (iCAS) recognized 32 predictive analytics professionals as the first recipients of its Certified Specialist in Predictive Analytics (CSPA) credential during the 2017 CAS Ratemaking and Product Management Seminar in San Diego.

The CAS Institute is a CAS subsidiary that offers credentials and educational opportunities for professionals working in highly specialized quantitative practice areas. CSPA credential holders possess practical knowledge of applied predictive analytics and data science used in data-intensive industry sectors.

For a number of years, new CSPA Susan Poole, FCAS, MAAA, has seen the expansion of predictive analytics in the insurance industry. "The CSPA credential combines a solid insurance foundation with predictive analytics to allow the practitioner to effectively tackle insurance-specific challenges," said Poole, a data scientist at SECURA Insurance Companies. "Attaining the CSPA credential has helped me to tailor my career path to incorporate an emphasis on predictive analytics," she said.

From the moment he first learned of the CSPA designation, Ron Lettofsky, ACAS, knew that it was something he wanted to pursue. Lettofsky, a newly credentialed CSPA, is a senior actuarial manager of claims analytics at Allianz Global Corporate & Specialty. "People who see the CSPA designation will know that I also have proven skills in predictive analytics and data management," he said.

The CAS Institute is accepting applications for the CSPA credential from experienced practitioners through November 30, 2017. For more information about the CSPA education program and the experienced practitioner application process, visit the iCAS website at thecasinstitute.org.

Kate Niswander is the marketing and communications manager for the CAS.

Certified Specialists in Predictive Analytics Recognized in March 2017

Seated, left to right: *Louise Francis, Susan Poole, Guangjin (Jim) Xiao, Stephen Stone,* **iCAS Leadership Advisory Council Chair Robert Miccolis**, *Ravi Kumar, Todd Lehmann and Cheng-Sheng Peter Wu.*
Standing, left to right: CAS President-Elect Brian Brown, *Christopher Monsour, William Frierson, Jeffrey Kinsey, Hernan Medina, Trent Goughnour, Gregory Hayward and Andrew Sutcliffe. Photo credit: Crown City Photography.*

The CAS Institute and its Community of Practice

The CAS Institute held its first-ever Community of Practice Event on March 27, 2017, in San Diego.

Designed to bring together advanced practitioners in predictive analytics and data science, the one-day event featured sessions on machine learning, external data, model design and deployment, ethics and risk governance.

Discussions at the event were led by distinguished practitioners in the disciplines of predictive analytics and data science, many of whom serve as subject matter experts for The CAS Institute.

In addition to the educational sessions, participants had opportunities to network and connect with others in the field.

Peter T. Bothwell speaks on a panel concerning ethics and risk governance at the Community of Practice Event. Bothwell is vice president, data science for The Hartford and a member of the iCAS Leadership Advisory Council.

Keep Current: Join iCAS

For notices about future events for predictive analytics professionals hosted by The CAS Institute (iCAS), become a member of iCAS at TheCASInstitute.org/membership. Dues are waived through September 2017.

A pioneer in data-mining, Louise Francis lends her expertise to a Community of Practice Event panel on external data. Francis serves as an iCAS subject matter expert and is president of Francis Analytics.

Participants of The CAS Institute's Community of Practice Event, held in San Diego on March 27.

How Can Insurers Find Real Value in Their Predictive Models?

BY CLAUDINE MODLIN, FCAS, MAAA

Experiment thoughtfully with practical implementation top of mind

There's a lot of conversation about new modeling approaches and novel sources of data poised to revolutionize insurance. This extraordinary industry transformation actually began about a decade ago. Analytical methods such as generalized linear models (GLMs) and decision trees were combined with new data sources, including credit attributes and prior insurance history, to improve pricing and underwriting sophistication. More recent developments, including vastly improved technology (e.g., hyper-scale computing and distributed storage), and an influx of new talent and availability of open-source programming languages and libraries, are providing even greater opportunities to explore what insights can be extracted from an increasingly wide array of data sources and formats. Are these influences triggering a revolution or evolution in insurance analytics? And how can insurers find real value in their predictive models?

Revolution or evolution? You decide.

Much of the buzz in insurance analytics circles is centered on investigating new analytical methods. Some of the techniques that are getting the most attention right now include gradient boosting machines (GBMs), penalized regression methods, neural networks, genetic algorithms and ensembles of different methods (Figure 1). While these methods are quite exciting, it's equally important for insurers to recognize the potential impact of new data sources. Including more diverse yet relevant data assets to an analysis adds far more predictive power than using more complex algorithms on existing data, as evidenced by usage-based auto insurance.

Additionally, insurers need to explore what types of problems different methods can address. No single method is perfectly suited to every business problem, and a variety of methods can add value at different stages of the modeling process. For example, topic modeling can help create new data features from unstructured text such as claims adjuster notes. Elastic nets can be useful in selecting factors for consideration in modeling. GBMs can help detect higher order interactions, and multivariate adaptive regression splines can help identify model hierarchies that capture complexity via a greater number of simpler models on well-defined segments. The end result is a more robust analysis. In fact, many interviews with Kaggle competition winners suggest that they do not necessarily credit their successes to the primary modeling method, but rather, to methods that enable better model inputs or corrections to the primary methods.

Figure 1

P&C insurers use a variety of modeling techniques
Non-GLM modeling methods are expected to gain momentum

	Now*	Two years
Generalized linear models (GLMs)	96%	100%
Decision trees	43%	63%
Model combining methods	37%	57%
Machine learning	17%	56%
Penalized regression methods	22%	46%
Random forest	19%	41%
Other ensemble methods	19%	41%
Gradient boosting machines (GBMs)	24%	39%

*Survey fielded September 7 – October 24, 2016.

Source: Willis Towers Watson 2016 U.S. Predictive Modeling and Big Data Survey

The inevitable question from the top: Where's the value?

As insurance company management hears more about advanced analytical methods, it begs the question of how these new methods really add value — or more specifically, how you even measure value.

To provide a meaningful answer for management, the analytics team should examine both statistical and financial value measures. Statistical measures, such as the Gini coefficient or Mean Absolute Error (MAE), have meaning among actuaries and data scientists but often don't provide management with an intuitive sense for value added. Moreover, the measures themselves don't often agree when ranking the accuracy of various methods. Financial measures are imperative for getting buy-in and gaining confidence from management. For example, when exploring new methods or new data for pricing and underwriting, estimating the loss ratio on actual out-of-sample claims can more effectively engage company management. We work with companies to design the right financial measures, including sensible underlying assumptions, to provide forecasts that make sense. In fact, in areas of the insurance company where data-driven solutions are relatively new, it's even more important to prove the financial value of the models to leadership.

Figure 2

Need help unlocking your analytical potential?

Willis Towers Watson offers advice to hundreds of P&C insurers globally, including carriers of different sizes that write many products and operate through different distribution channels. We pioneered the use of GLMs in pricing, and continue to innovate, harnessing new techniques to meet new challenges. We help companies assess the suitability of methods across a variety of dimensions, including not only predictive power but interpretability, ease of implementation, relative effort and execution speed. Methods such as GLMs are well-accepted in areas such as pricing because of their transparency, ease of implementation (in traditional table-based rating engines) and execution speed. Other insurance applications place different values on the various dimensions. For example, producing direct mailing lists based on expected profitability and likelihood to buy does not require high levels of transparency, and implementation requires a list of addresses rather than inputs to table-based engines.

We help companies explore and find value in new data, methods and applications in a variety of ways:

- Evaluate new data assets.
- Train client teams in machine learning techniques for a defined problem of choice.
- Deploy machine learning techniques to sharpen existing (traditional) models.
- Assist with machine learning in applications that may not require high transparency (e.g., topic modeling adjuster notes to create new structured fields, and examining voice data for opportunities in improved customer satisfaction).
- Streamline modeling processes and introduce hierarchies.

Software that addresses the entire pricing workflow

Willis Towers Watson's trusted pricing software, used by many of the world's largest insurance groups, can support your entire pricing workflow, including deployment. Radar Base, which is used to assess and compare model results and perform dynamic impact analysis on real customer data, can now import a variety of model forms built in other programming environments.

Adding to the Radar platform, Radar Live provides a single, holistic environment for analytics and deployment, undiluted by systems constraints (Figure 2). Radar Live is more than an external rating engine. It enables a wide range of analytics to be deployed in real time at point of sale — from traditional rating structures to complex pricing algorithms with sophisticated embedded risk models. Any risk classification, rule, model or calculation programmed in Radar Base can be uploaded into the Radar Live production environment via a preproduction and testing stage. This not only provides great pricing flexibility and responsiveness to market developments but also creates material operational efficiencies and reduces the risk of costly errors in programming rates in multiple environments.

What's needed to change?

Analytics are transforming the insurance industry. However, this requires thoughtful experimentation and constant consideration of implementation requirements.

For more information, email claudine.modlin@willistowerswatson.com.

Predictive Analytics: What's Next?

BY ROOSEVELT C. MOSLEY JR., FCAS, MAAA, CSPA

A few years ago I delivered a presentation entitled, "Beyond the Credit Score." By 2010, the use of credit-based insurance scores in personal lines insurance had become standard, yet these scores were still subject to significant regulatory and consumer scrutiny. As a result, many insurance companies began asking questions about alternatives to credit in an effort to develop a plan to move beyond the use of credit score if it ultimately became necessary.

My presentation answered insurers' questions as to why the use of credit scores was so successful. The three primary reasons were:

- Credit scores provided significant separation of indicated risk differences.
- There was a reasonable distribution of insured risks across the credit-score scale.
- There was not a significant overlap of credit score with existing risk characteristics.

Then, using these three criteria, I identified the following areas that would provide benefits similar to those observed through the use of credit scores:

- Usage-based insurance (UBI).
- More refined territory definitions.
- More descriptive insured property information (vehicle characteristics, property characteristics, etc.).

Since 2010, significant progress has been made in each of these areas.

But one advancement I discussed has not moved as quickly — the use of predictive modeling techniques beyond Generalized Linear Modeling (GLM) in the development of rating plans.

Insurance companies began to use GLMs in rating plan development in the late 1990s and early 2000s. GLMs were a significant advancement over older techniques as they allowed companies to consider the impact of all factors at once, thus removing the distributional bias from the indication process. This provided a more accurate representation of the impact of each risk characteristic on loss costs.

However, GLMs have their limits. First, GLMs are linear models, despite being generalized. While this linear assumption is generally reasonable for insurance data, it tends to be less accurate at the extremes. GLMs tend to underestimate the risk potential of policyholders with the lowest expected loss costs, while they overestimate it for those with the highest.

GLMs also assume that the risk associated with a combination of factors is represented purely by the product of the risks associated with each individual underlying factor. For example, the initial assumption built into a GLM for an auto risk is that the percentage increase in expected loss cost for a driver with a prior accident is the same regardless of whether the driver is 17 or 47 years old. This concern can be addressed by the use of interactions, but higher-order interactions are difficult to incorporate into a GLM, and including a full interaction is overkill if you are only interested in its significant portions.

In a world where insurers are looking for "what's next" … non-GLM techniques should be one of the answers.

The use of other modeling techniques allows companies to address these issues and find significant lift in their rating plans. These methods include, but are not limited to, Decision Trees, Neural Networks and Gradient Boosting. Applying these approaches to supplement the power of a GLM yields a more predictive result than can be obtained from either independently.

In analyses including non-GLM techniques, we have been able to consistently achieve results showing a range of indicated relativities of at least 3 to 1. This additional lift was identified over and above what the GLM was able to achieve. This indicated lift rivals that of credit score and is achieved simply by using the information already being considered in a rating plan.

In a world where insurers are looking for "what's next" in order to gain or maintain a competitive advantage, non-GLM techniques should be one of the answers. Exploration of these approaches can provide insurance companies with a significant competitive advantage. In this case, the next significant move forward in rating could actually come from within.

Roosevelt Mosley is a principal and consulting actuary with Pinnacle Actuarial Resources, Inc.

CONGRATULATIONS

TO THE RECIPIENTS OF THE CERTIFIED SPECIALIST IN PREDICTIVE ANALYTICS CREDENTIAL

The CAS Institute recently awarded the Certified Specialist in Predictive Analytics (CSPA) credential to the following 32 individuals:

Avraham Adler, CSPA, FCAS, MAAA, CERA — Guy Carpenter & Co. LLC

Joel Atkins, CSPA, FCAS, CPCU — CNA Insurance Companies

Shane Barnes, CSPA, FCAS — The Hartford

Andrew Brown, CSPA — Guide One Insurance Group

Richard Crabb, CSPA, FCAS — University of Wisconsin - Madison

Denise Christophel, CSPA, CPCU — Sentry Insurance

Linhui Dong, CSPA — Munich Re America

Louise Francis, CSPA, FCAS, MAAA — Francis Analytics & Actuarial Data Mining Inc.

William Frierson, CSPA — Willis Towers Watson

Luyang Fu, CSPA, FCAS — The Cincinnati Insurance Companies

Trent Goughnour, CSPA — Pinnacle Actuarial Resources, Inc.

James Guszcza, CSPA, Ph.D., FCAS — Deloitte Consulting, LLC

Gregory Hayward, CSPA, FCAS, MAAA, FCIA, CERA — State Farm

Jeffrey Kinsey, CSPA, FCAS, MAAA — State Farm

Ravi Kumar, CSPA, ACAS, MAAA — QBE North America

Todd Lehmann, CSPA, FCAS, MAAA — Quincy Mutual Fire Insurance Co.

Ronald Lettofsky, CSPA, ACAS — Allianz Global

Weiting Lu, CSPA — Oliver Wyman Actuarial Consulting

Zachary Martin, CSPA, FCAS, FSA, MAAA — Zurich North America

Hernan Medina, CSPA, CPCU — ISO

Stephen Mildenhall, CSPA, Ph.D., FCAS, ASA, MAAA, CERA — St. John's University

Christopher Monsour, CSPA, FCAS, MAAA — CNA Insurance Companies

Roosevelt Mosley, CSPA, FCAS — Pinnacle Actuarial Resources, Inc.

Susan Poole, CSPA, FCAS, MAAA — SECURA Insurance Companies

Ernesto Schirmacher, CSPA — Liberty Mutual Insurance

Stephen Stone, CSPA, FSA — Agam Capital Management

Andrew Sutcliffe, CSPA — Allianz Global

Rebecca Vessenes, CSPA, Ph.D., ASA — Liberty Mutual Insurance

Cheng-Sheng Peter Wu, CSPA, FCAS, ASA, MAAA — Deloitte Consulting, LLP

Guangjin Xiao, CSPA, FCAS, MAAA — CNA Insurance Companies

Jonathan Zabek, CSPA, MSPA — Franklin Mutual Insurance Company

Ya Zhang, CSPA — One Beacon Insurance Group

Added Values: Breathe Life into P&C Projections

BY STEPHEN URBROCK

Lifetime value (LTV) style calculations may have made their name in life insurance but are now proving their worth to property and casualty (P&C) businesses. Neil Covington, director of solutions management for FIS's P&C business, explains the lure of LTVs — and how predictive analytics can extract even more value from P&C projections.

Why LTVs aren't just for life

Life insurance, term assurance and mortgage contracts typically span decades. So, it has traditionally made sense for life insurers to assign an LTV to their customers — and project the long-term, total value each customer or contract will represent.

P&C policies, by contrast, tend to last no more than a year. But in the first year, the upfront cost of selling a new policy may take a significant portion of income, as companies try more innovative ways to maximize policy retention. This "new business strain" may even exceed year-one profit margins, meaning the more new policies you sell, the bigger your loss over the year.

The answer for many P&C firms has been to look past the first-year accounting period to the income that a policy could bring in over a lifetime — its LTV. From year two onwards, income and margins will soon overtake the initial outgoings — and the longer the policy is renewed, the greater the LTV.

Value your customers

The LTV comes into its own when it is used to reflect the value of not just individual policies but also customers. If a customer has taken out home and auto policies with the same insurer, each contract will carry its own LTV. Added together they will reveal the total value of that customer's relationship with the company and help build a holistic view of their value beyond the balance sheet. This aggregated view will come backed with a wealth of policy rating data that is ripe for predictive analytics.

Empower your projections

Online sales channels, telematics technology and increasingly digital operations make it easier for today's insurers to gather behavioral information on their customers. Predictive analytics can extract more meaning from rating data and use information from all of a customer's policies and other products to forecast future value.

Why, for example, is one customer's LTV higher than another? By applying predictive analytics techniques, you can drill down into the complex combination of factors involved — from age and location to lifestyle choices — and identify which customer segments to target with which products.

As well as informing new business marketing strategies, this approach can help you retain and cross-sell more effectively to existing customers. Given his or her profile, what are the chances of a customer renewing a policy or extending their cover?

> **"We've certainly seen growing interest from P&C insurers in LTV analysis — and predictive analytics systems are becoming an important part of their risk management toolkit."**
> **—Derek Chapman, Principal, Merlinos & Associates, Inc.**

To answer such questions and build predictive models for LTVs, you can apply the same generalized linear modeling (GLM) techniques often used for pricing. You can also show how an LTV may evolve in the future and the best ways to improve or protect it. And with machine learning, a growing capability of advanced analytics systems, the accuracy (and value) of these projections will only improve over time.

Gain a platform for growth

Key to putting projections into practice will be a powerful integrated risk management platform that can support full capital modeling projections alongside individual LTV and customer value calculations. With built-in predictive analytics tools, it will need to handle both GLM and clustering analysis to derive full value from data.

Investing in a solution of this kind will soon pay dividends, by helping you better understand the dynamics of your business and its risks. Above all, it should give you the tools to help better meet the needs of your customers — and improve shareholder value and returns.

For information please contact:
Stephen Urbrock
FIS Insurance Software
Cell: 404.205.9156
Email: stephen.urbrock@fisglobal.com
www.prophet-web.com
www.fisglobal.com

Predictive Analytics Providers Directory

Organizations providing predictive analytics products and services.

A.M. Best
Douglas Hamadyk
908-439-2200 x5753
www.ambest.com

Actuarial Resources Corporation
Chris Peek
913-451-0044
www.arcval.com

CBIG Consulting
Jim Grosspietsch
800-334-2078
services@cbigconsulting.com
www.cbigconsulting.com

CGI
Kris Komassa
512-791-7328
www.cgi.com

Conning
Lorraine Hritcko
860-299-2403
lorraine.hritcko.com
www.conning.com/products/risk-management

CoreLogic
Stephanie T. Grayson
877-849-1023
www.corelogic.com

DataRobot
Satadru Sengupta
617-301-2471
satadru@datarobot.com
www.datarobot.com/insurance

Decision Research Corporation
Rick Young
800-836-6057
www.decisionresearch.com

Digital Recognition Network
Amanda Kirk
817-710-7789
www.dmdata.com

Earnix
Adi Bar-Lev
972-73-706-7247
adib@earnix.com
www.earnix.com

Easy2Comply
David Leichner
800-429-4391
www.easy2comply.com

Ernst & Young
Gary T. Ciardiello
212-773-1377
gary.ciardiello@ey.com
www.ey.com

FinCad
Lori Bryenton
604-957-1216
www.fincad.com

FIS Insurance Software
Stephen Urbrock
404-205-9156
stephen.urbrock@fisglobal.com
www.prophet-web.com

Gross Consulting
Chris Gross
651-293-8008
chris.gross@cgconsult.com
www.cgconsult.com

Guidewire Software
Tom Kasel
651-470-0748
info@guidewire.com
www.guidewire.com

IBM Algorithmics
Curt Burmeister
914-499-1900
www.us.ibm.com

IHA Consultants, Inc.
Mark Zanecki
919-260-3291
mark.zanecki@ihaconsultants.com
www.ihaconsultants.com

Insight Decision Solutions Inc.
Claudia Wetzel
416-479-0384
info@insightdecision.com
www.insightdecision.com

Insureware Pty Ltd
Adi Kedmi or Sylvia Gooch
61-3-9533-6333
admin@insureware.com
www.insureware.com

KPMG
James Christou
917-756-0589
www.kpmg.com

LexisNexis Risk Solutions
Clare Louise Southcombe
561-212-7375
clare.southcombe@lexisnexis.com
www.lexisnexis.com

Milliman
Brian Brown, FCAS, MAAA
312-499-5660
brian.brown@milliman.com
www.milliman.com

Moody's
Lindsay Hagans
415-874-6350
www.moodys.com

Nexus Risk Management
Brianne Krysiak
312-857-4401
www.nexusrisk.com

Numerix
212-302-2220
marketing@numerix.com
www.numerix.com

Octo Telematics North America
Nino Tarantino
617-916-1080
info@octousa.com
www.octousa.com

Pinnacle Actuarial Resources, Inc.
Julie Calmès
309-807-2300
jcalmes@pinnacleactuaries.com
www.pinnacleactuaries.com

PolySystems Inc.
Bob Keating
312-332-5670
www.polysystems.com

PricewaterhouseCoopers
Lisa Slotznick
646-471-4000
www.pwc.com

PRMIA
Janet Tritch
612-216-4017
janet.tritch@prmia.org
www.prmia.com

Red Mountain Technologies
Chris Whipple
205-414-2721
www.redmountaintech.com

Reserve Prism
Hai You
608-239-1670
hyou@reserveprism.com
www.reserveprism.com

Risk Lighthouse LLC
Jasmine Speights
678-732-9112
www.risklighthouse.com

rPM3 Solutions, LLC
Rob Eckels
410-384-9491
reckels@rpm3solutions.com
www.rpm3solutions.com

Salford Systems
Amy Baldwin
619-543-8880
www.salford-systems.com

SAP
Andrew Winick
610-661-7519
www.sap.com

SAS
919-677-8000
www.sas.com

SimErgy
Sim Segal
646-862-6134
sim@simergy.com
www.simergy.com

Ultimate Risk Solutions
Isabella Dumont
845-825-1494
www.Ultirisk.com

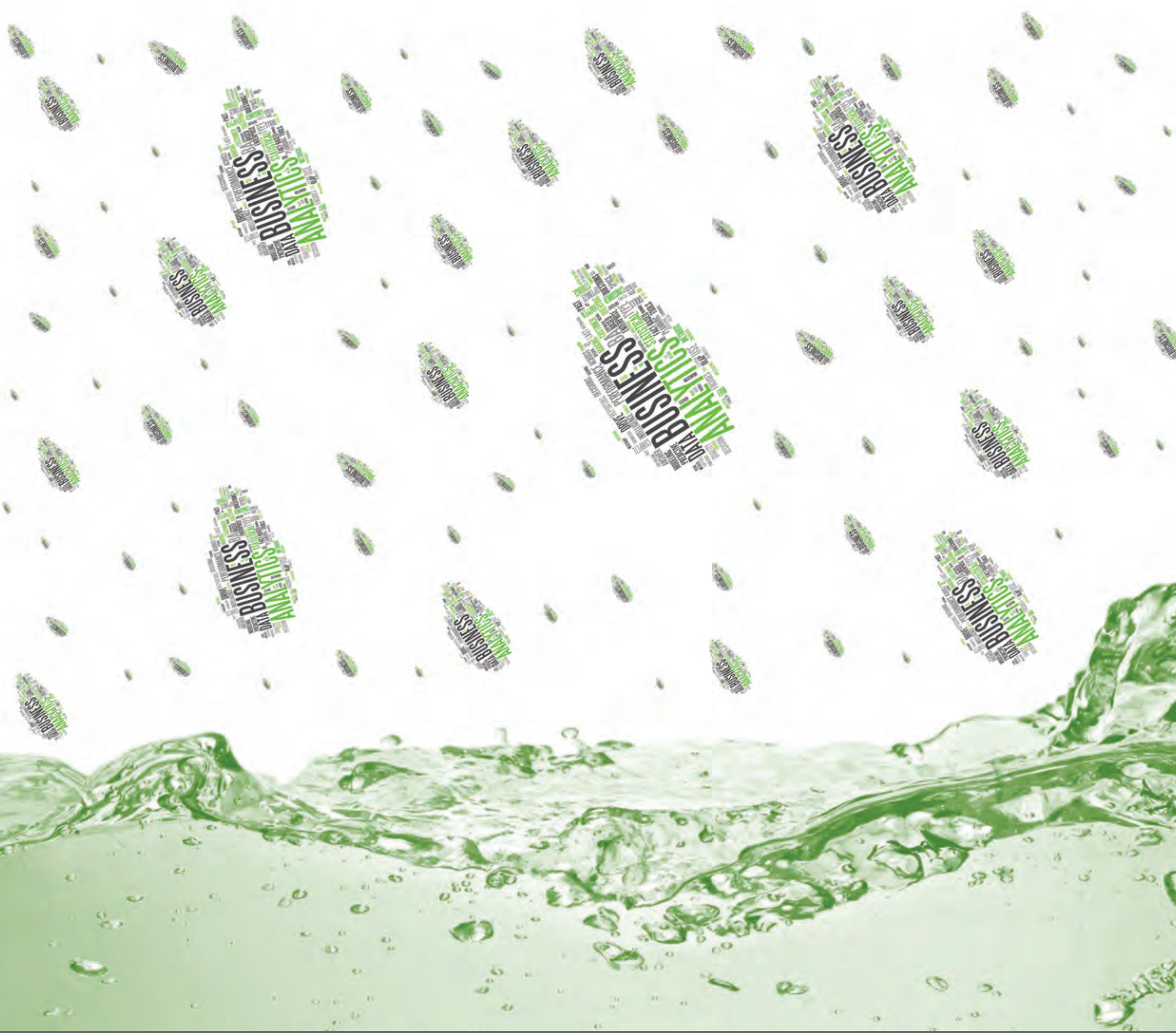
Valen Technologies
Kirsten Marr
800-280-3304 x235
www.valen.com

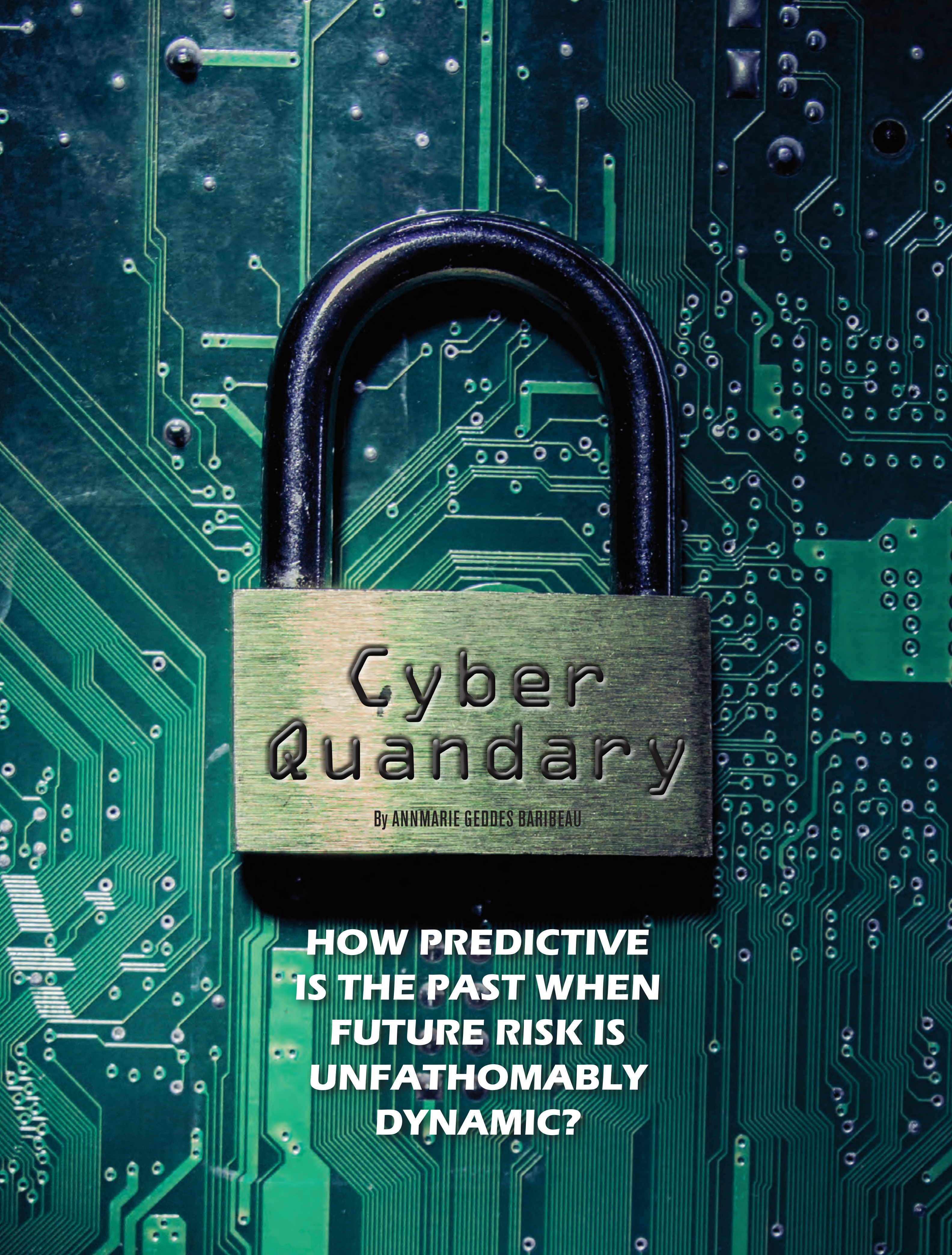
Verisk Analytics/ISO
Chip Chaffee
201-469-2394
www.verisk.com

Willis Towers Watson
Claudine Modlin
805-499-2164
claudine.modlin@willistowerswatson.com
www.willistowerswatson.com

Wolfram Research
Michael Gamman
217-398-0700
www.wolfram.com

Wystar
800-505-9076
rsmarketing@wystar.com
www.wystar.com

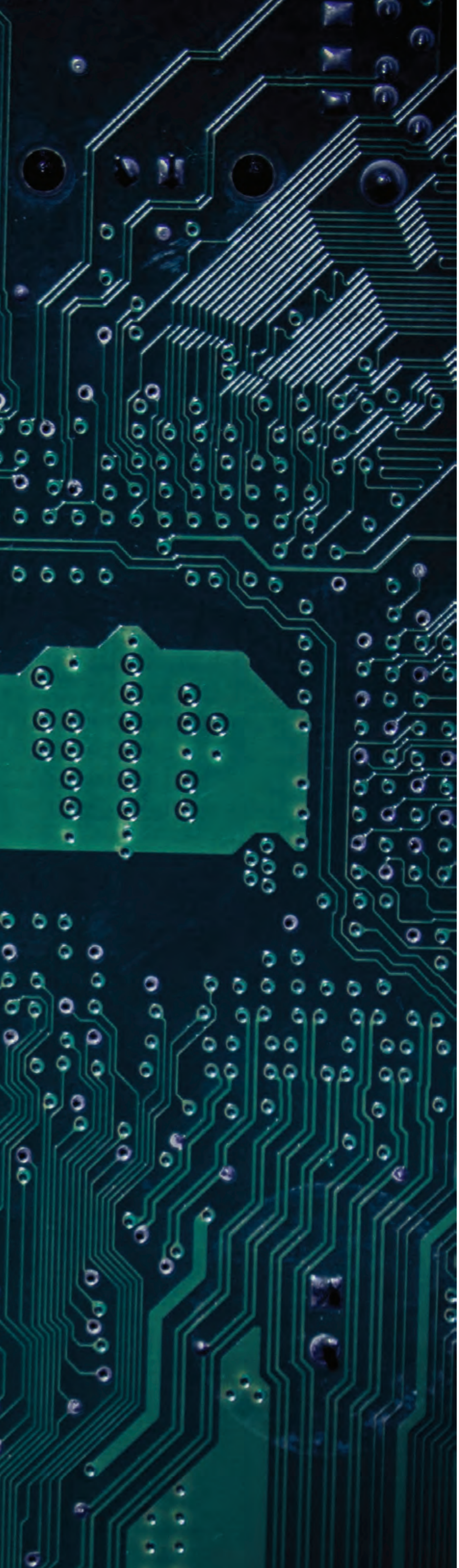

BUSINESS
ANALYTICS
DATA



Cyber Quandary

By ANNMARIE GEDDES BARIBEAU

HOW PREDICTIVE IS THE PAST WHEN FUTURE RISK IS UNFATHOMABLY DYNAMIC?



Cyber insurance actuaries face a challenging reality. While claims data availability is growing, relying on past information to predict future losses is just not enough. As the burgeoning cyber insurance market, emerging cyber vulnerabilities and unfathomable risk continue, actuaries are charting new territories to help insurers confidently write coverage.

To appreciate the pace of change, consider how the cyber world has shifted in the past three years. Harrowing data breach headlines began to dramatically boost cyber insurance sales. Public awareness of the internet of things was just beginning, and cloud computing was considered safe.

Fast forward to today. Increasing cyber insurance sales have led to additional claims data, but risk and coverage continue to change. Ransomware claims are on the rise while greater connectivity from cloud computing, the internet of things and sophisticated automation are introducing new cyber vulnerabilities. “The attack surface of an organization is loosely the sum total of all points of vulnerability,” says Jon Laux, head of cyber analytics at Aon Benfield.

More organizations are buying cyber insurance for the first time or are expanding it. While the insurance industry has responded with higher limits, customers also want cyber insurance to cover more types of risk.

Market Overview

Cyber insurance has existed in some form for about 20 years, and it remains a developing insurance line. Cyber insurance is offered through stand-alone policies or as an add-on to traditional business coverage such as general liability. There are more than 60 different cyber insurance carriers that offer great variation in coverage scope, policy triggers, definitions and exclusions, according to the Aon survey “Cyber — The Fast Moving Target,” released in June 2016.

Coverage categories identified by Marsh & McLennan Companies include business income, data asset protection, event management, cyber extortion, privacy liability, network security liability, privacy regulatory defense costs and media liability. These competing policy forms have little uniformity, says Robert Parisi, Marsh’s cyberrisk product leader.

Growth as measured by premium continues in the United States. Estimates of growth vary. Parisi points out that the general market consensus is that 2016 closed out just shy of $3 billion in gross written premium, up from $2 billion in 2015 and $1.5 billion in 2014.

Aon’s estimate is lower, with the current size of the global cyber insurance market at about $2.5 billion. Laux says that the U.S. market is 80 to 90 percent of the global market and that 60 to 70 percent are stand-alone policies. Aon estimates the global market is expected to grow to $7.5 billion to $10 billion by 2020.

Stand-alone cyber insurance purchases among Marsh clients in the U.S. rose 27 percent from 2014 to 2015, according to Marsh’s *MMC Cyber Handbook 2016*, released in November 2016. This was mainly driven by an increasing awareness and appreciation of cyberrisk, particularly at the boardroom level. “We saw the same increase (26 percent growth) between 2015 and 2016,” Parisi says.

Despite the unfathomability of

future risk, insurers continue to compete for more business. Generally, cyber insurance has been profitable.

Michael Solomon, a consulting actuary with The Actuarial Advantage, Inc., calculated that cyber insurance stand-alone polices in the United States have a loss and loss adjustment expense (LAE) ratio of 65.2 percent based on figures from the National Association of Insurance Commissioners' (NAIC) *Cybersecurity and Identity Theft Insurance Coverage Supplement* for 2015 statutory filings. Using updated NAIC information, Aon estimates the loss and LAE ratio for the United States at 50 percent for stand-alone coverage and 41.5 percent for both stand-alone and cyber coverage packaged with other insurance lines.

"Profitability has varied widely, with even carriers who have substantial books of cyber insurance business recording loss and LAE ratios from as low as one percent to as high as over 160 percent," says Alex Krutov, president of Navigation Advisors LLC. Krutov, who started and was the first chair of the CAS Cyber Insurance Task Force, says that the median loss and LAE ratio is less than 40 percent for narrowly-defined stand-alone cyber insurance.

Companies buy cyber coverage for various reasons. According to the June 2016 Aon survey, the majority of survey respondents that purchase cyber insurance (68 percent) cite balance sheet protection as the main motivator.

Customers also want broader coverage. The growing need to cover business interruption and contingent business interruption risks, for example, is due to expanding dependence on technology, Parisi observes. "(This is) the most significant thing going on in the cyber market," he explains, because traditional business interruption coverage generally does not cover losses when cyber incidents are the cause.

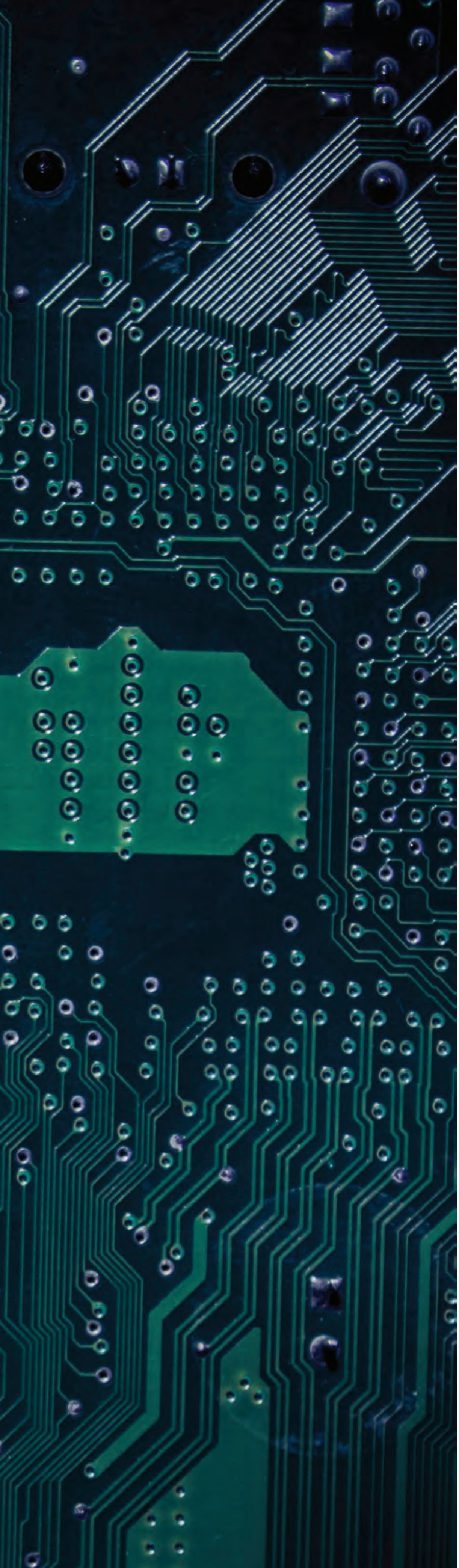
Business interruption, both during a system breach and post breach, was rated as the top cyberrisk concern, according to the Aon report. Bodily injury/property damage, which is generally covered in first- and third-party coverage, was the lowest rated concern of respondents.

"The reality is the traditional cyber insurance product in the U.S. has largely been developed to address data breaches with some ancillary coverage," Laux says. "That leaves a lot of ground uncovered for a manufacturing company that could have major cyberrisk if hackers get into the control systems of a manufacturing plant."

Policies are starting to pick up business interruption risk, Parisi observes. "What we have seen in cyber (insurance) is a fairly quick research and development cycle with awareness of the risk being recognized by the buyer and then carriers assessing how much of that risk they are willing to accept."

More unintended consequences from the internet of things, for example, will lead insurers to ask more about their own risk aggregation and possibly to change how they provide coverage for risks associated with connected devices, according to Marsh's report, "The US Casualty Market In 2017: Our Top 10 List," published in January. Due to greater connectivity, the report notes, "[t]he boundaries blur between traditional product

Despite the unfathomability of future risk, insurers continue to compete for more business. Generally, cyber insurance has been profitable.



liability and cyber insurance."

Data Desire

While all indications point to a growing demand for cyber insurance, insurers face the ongoing quandary of providing coverage amid a plethora of unknowns.

Cyber insurance is different from other types of insurance in at least one significant way: The traditional actuarial adage that the past is the predictor of the future has limited application. Further, actuaries involved in cyber insurance differ on whether there is enough quality data for reliable modeling.

"Many people are still saying we just need more years of data," Krutov says. "However, in cyber insurance, you cannot just use historical data the way it is done in traditional actuarial models — even if we had a lot more of this data."

"Actuaries and insurers have enough claims data to tell a useful part of the story but not the whole story," observes Michael Solomon. He sees the lack of data as an opportunity for actuaries to demonstrate their value, which is evaluating risk using their experience and unique actuarial judgment.

While claims data offers more insight than insurers had three years ago, the data situation remains less than ideal. Sources agree that data quality is a pressing concern. Data standardization "will go a long way" to improve data quality and consistency, says Robert Hartwig, clinical associate professor and co-director of the University of South Carolina's Center for Risk and Uncertainty Management. Hartwig is the co-author of the Insurance Information Institute's report, "Cyberrisk: Threat and Opportunity," released in October 2016.

Hartwig says that cyberrisk has the potential to permeate every insurance line. The internet of things, for example, will affect auto and home insurance. Advanced medical devices will have an impact on workers' compensation. "There will even be workers who will be wearing, ingesting, or implanting devices to assess workplace injuries," he says.

Since risk profiles are growing in relevance, insurers need to collect more information about their insureds. "When we ask insurers to provide us with information to run models, the reality is that most insurers simply are not capturing enough data," Laux says. A set of common core data requirements for cyberrisks was published in 2016 by Lloyd's of London in conjunction with modeling firms Risk Management Solutions and AIR Worldwide. This was a good start, Laux notes, but it will be some time before insurers are actually capturing the relevant data fields for modeling.

The cyber insurance underwriting process evolved out of errors and omissions (E&O) coverage questionnaires. While cyber insurance has things in common with E&O insurance, Laux says, the information needed to thoroughly grasp cyberrisks goes well beyond what underwriters are typically asking. "The information needed is much more subtle," he adds.

Data collection, however, relies greatly on the willingness of underwriters to ask more questions. "A real issue is not wanting to impede the sale," Parisi says. Some insurers are looking at organizations in specific industries with so much scrutiny that Parisi observes that they risk "paralysis by analysis." This is especially true for business segments hit hard by breaches over the past two years and for new coverages, such as contingent business interruption coverage.

Underwriters also continue to have tremendous influence on what insurance companies ultimately write. Actuarial influence varies widely depending on the insurer.

Alternative Information

Since claims data has limited use in predicting losses amid the continual influx of emerging cyber vulnerabilities and losses, actuaries are seeking alternative data sources and methods to prepare insurers to cover future claims.

Actuaries need to loosen their expectations about what data should look like and where to find it, Laux says. Because each data set has strengths and weaknesses, the key is to use the right data to solve the right part of the problem within tight parameters.

Since insurers and cybersecurity firms share an interest in understanding the factors that contribute to cyberrisk, along with frequency and severity of incidents, experts agree that using nontraditional insurance data makes sense. "The reality is there is a massive amount of data being gathered every day, but it sits in the systems that defend companies, [in] technology servers and on the internet for those who know how to look for it and capture it," Laux says.

"The most interesting area we're grappling with as an industry," he says, "is how to take all this technology data, which can be seen as leading indicators or signals of cyber compromise, and turn that into intelligence about the probability of a claim and its size."

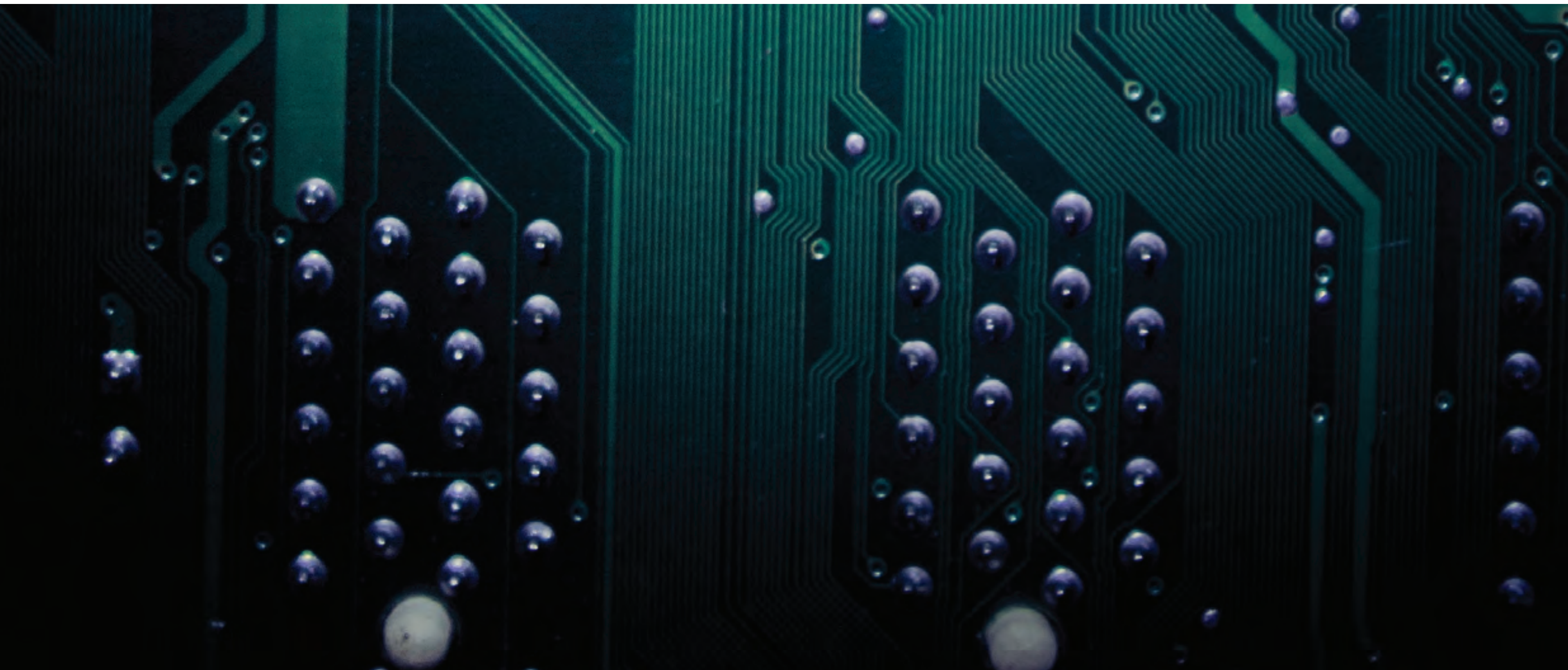
Incident data from cybersecurity firms and other sources, Krutov says, offers insight into both frequency and severity of potential cyber events. Information on uninsured financial losses is another nontraditional data source, he suggests. "Actuaries are usually very reluctant to use this type of data because there is a big difference between insured and uninsured financial losses even where loss events appear to be similar." While the reluctance is justified, Krutov believes this information is useful in modeling risk for both existing and new kinds of cyber insurance coverage.

Before his company had sufficient claims data from 10 years of selling cyber coverage, it was still able to develop models using other sources of data, says Adam Rich, actuary and head of specialty lines analytics for the Beazley Group. "We simply built a model to try and parameterize the things that might give rise to payments under a policy," he adds. Such factors include claims frequency by size, amount of forensic or legal expenses, the cost of notification services and monitoring of credit — as well as the chance that those services would be needed or requested.

Solomon recommends keeping up with cyber reports and paying attention to international trends — especially in Western Europe — since cyber incidents are independent of geographic location. "We are looking for causes of loss and where the malicious actors are going, which can be learned from the experience of other countries," he says.

Modeling Uncertainty

Experts have different opinions concerning just how far along cyber modeling is developing.



[Hartwig] likens the state of cyber insurance modeling to the late 1600s, when Lloyd's of London was insuring ships headed to the New World with little information for determining risk potential.

Poor quality data requires insurers to analyze cyberrisk "from a technical rather than a statistical point of view," according to The Geneva Association report, "Ten Key Questions on Cyber Risk and Cyber Risk Insurance," published in November 2016. When it comes to factoring in the fast-changing risk landscape, the report states, "There is no established method to model cyberrisk, and not much research has been done so far."

Hartwig says that cyber modeling — especially for catastrophic loss — is in its infancy. He likens the state of cyber insurance modeling to the late 1600s, when Lloyd's of London was insuring ships headed to the New World with little information for determining risk potential. "We actually have zero data on the much-feared 'cyber Pearl Harbor' or 'cyber 9-11' attack scenarios," he says.

Laux contends that model development has come a long way in just the past couple of years. He explains that, in 2015, the discussion at industry conferences was about whether cyberrisk could be modeled. Now, the conversation has evolved from abstract ideas to a working discussion. "And I expect the pace of development will continue," he adds.

Modeling innovation is largely coming from cybersecurity firms that use analytics to protect their customers, Laux says. "The single biggest area of excitement in terms of analytics in this space has to do with the influx of technical and cybersecurity firms providing analytics," he observes. "Insurance has become a use case."

Krutov offers that cybersecurity models tend to focus on quick detection of anomalies. This makes them less useful for insurance pricing because a longer-term time horizon of at least a year is necessary. "Very often the choice of models, whether complicated or of the back-of-the envelope variety, is driven by data availability," he says.

Rich insists there is enough data available to create pricing models. Beazley deploys several different cyber models according to the insured's characteristics and the amount and type of coverage. The insurer also monitors an insured's portfolio from an aggregate view while segmenting it for emerging trends.

To determine the likelihood that a customer could become a cyber incident target, Laux says that Aon recommends conducting a risk analysis of each customer using factors such as company size, industry and security posture. Cyberrisk modeling needs to go beyond considering previous cyber losses. It needs to anticipate the shifting nature of exposures and to contemplate cyberrisk as a peril rather than as a narrowly defined coverage.

For severity modeling, Solomon looks at costs from losses already covered by other types of insurance, such as intellectual property or theft. He uses this information to determine potential insurance losses due to cyber incidents. Outside data can also provide insight. "Even the cost of business interruption not from the insurance field is helpful," he adds.

According to The Geneva Association report, modeling frequency and severity can be accomplished by using extreme value theory and the peaks-over-threshold approach. "Heavy tail distributions have been proposed, [e.g.,] the power law or the log-normal distribution for the severity and negative binomial distribution for the frequency," the report notes.

Emerging Risks

Although inroads are being made to address data availability, modeling emerging risks will always be challenging. Continuous revisions of modeling techniques are necessary in a quickly-changing technological environment, the Geneva Association report notes.

Sources agree that anticipating the cost potential for emerging risks is challenging. Besides using cybersecurity data to detect risks, claims and underwriting professionals can also provide insight into emerging trends.

The growth of ransomware is one emerging trend. Criminals are figuring out that ransomware is a much more efficient way to make money, Solomon says.

At Beazley, claims stemming from ransomware attacks have more than quadrupled in 2016. Nearly half of these attacks occurred in the health care sector, according to an internal study posted on its website in January. Such attacks will double again in 2017, Beazley projects. "We are paying out more of those, and even our policies are starting to include language to explicitly state how much we will pay for ransomware," Rich states. The company pays in Bitcoin as a service to customers, he adds.

Two more emerging risks stem from greater connectivity through the internet of things and cloud computing, Parisi says. Both are making insurers and insureds cautious, Parisi notes.

"As more things get connected to the internet," Rich explains, "the distribution of connectivity might change an insured's exposure."

Conclusion

A convergence of conditions is having an influence on cyber insurance. Insurers will continue to face pressure to expand coverage amid unpredictable new risks stemming from greater automation and connectivity from cloud computing and the internet of things. The much-feared and impossible-to-predict "cyber catastrophe" that could cascade through multiple organizations remains an aggregation concern for insurance companies.

While data is becoming available for more meaningful models, underwriters continue to rely greatly on their own judgment for writing coverage. Just as it takes time for underwriters in other lines to trust modeling results, so too will actuarial influence grow stronger.

Cyber coverage policies and data collection call for greater standardization — especially as small- and medium-sized organizations are beginning to realize their own vulnerabilities and need for cyber insurance. Hartwig predicts there will be robust penetration in the middle markets in 10 years as the insurance industry grows more comfortable covering cyber "fender benders."

Hartwig says that the actuarial profession will need to update exams and provide continuing education so its practitioners will remain relevant. He predicts, "An actuary specializing in cyber and technological risk will have a great future."

Annmarie Geddes Baribeau has been covering actuarial topics for more than 25 years. Her blog can be found at http://insurancecommunicators.com.

The much-feared and impossible-to-predict "cyber catastrophe" that could cascade through multiple organizations remains an aggregation concern for insurance companies.

Darwinian Theory Meets Insurance Analytics BY JIM LYNCH

"Survival of the Fittest" is a familiar phrase in biology, but actuaries at the Casualty Actuarial Society's Ratemaking and Product Management seminar in March got a lesson in how to let Darwin's theory improve insurance analytical practices.

Eliade Micu, FCAS, director of research and development at Clyde Analytics, gave attendees an overview of how to use the rules of biological natural selection to solve complex insurance problems in a session titled, "Genetic Algorithms with Applications in Insurance."

The two most effective problem solvers in nature are the human brain and natural selection, said Micu. Using brainpower to solve problems is, well, a no-brainer. Using evolutionary methods is a bit more revolutionary.

The approach — known as evolutionary computing — makes sense when modeling complex problems, Micu said, particularly when there are many solutions, or when you have a good solution to a problem but suspect that a radically different approach might yield a better one.

The concept stretches back to the 1940s; its earliest success was from devising timetables for university classes. The problem is large and complex, with a lot of constraints, including that:

- There are thousands of classes and thousands of classrooms in which to conduct classes.
- Students prefer to not have too many classes in a day, while faculty wants classes concentrated in as few days as possible, to leave more time for research.
- Classes early or late in the day have limited appeal.

Winning solutions typically tend to defy human expectation. A classic example is the design of a structure to attach a satellite dish to a space shuttle. Human designers came up with a standard rigid structure, but struggled to account for the fact that in space you don't have gravity and an atmosphere to thwart vibrations once they start.

The winning model vaguely resembles what humans designed, but really looks more like the skeleton of a beer can that has been squeezed in the middle. It appears semi-squished and twisted. It looks delicate as a wounded butterfly, but it works 20,000 percent better than the standard rigid structure.

Evolutionary computing methods "tend to produce strange-looking things that humans would not come up with," Micu said. "But they seem to work … just magical; they just work."

Natural selection works by having many individuals compete and allowing successive generations to evolve. Eventually a certain population stumbles upon a winning set of mutations and crowds out opponents.[1]

Eliade Micu, FCAS

Evolutionary computing works in a similar way. The programmer wants to solve a problem that might defy a conventional approach. So he sets up a model that mimics the problem and sets several competing solutions at the problem. And he gives them the opportunity to attack the problem anew in successive generations.

[1] Rodrigo C. Barros, Andre C. P. L. F. de Carvalho, Alex A. Freitas. *Automatic Design of Decision-Tree Induction Algorithms* (Cham, Switzerland: SpringerBriefs in Computer Science, 2015), 3.

Each new generation is created through a sort of algorithmic procreation. Solutions pair off in a semirational, semirandom fashion, and their offspring are a new solution.

And just like in real life, parents can pass along their traits, and those traits mutate randomly. Mutations happen the same way they do in nature. The equivalent of a parental gene can mutate and pass along a new trait, or there can be a crossover, in which a portion of one parents' genetic code swaps places with the code located at the same spot on the other parent.

And as in real life, some new solutions thrive; some do not. One way evolutionary computing differs from the real world is that the modeler needs to watch for characteristics that appear superior but are losing out because of bad luck.

The phenomenon is called genetic drift, and it is more likely to occur, both in nature and in computing, when the population size is small. Modelers want to avoid genetic drift because it is completely caused by chance and not by the superiority of a characteristic. Modelers want randomness to create winning designs, not eliminate them.

With genetic drift, Micu said, "Highly fit individuals may be lost."

There are other design challenges.

- What characteristics do you give the first generation? (You can assign them randomly, Micu said, but it isn't unusual to seed the population with known solutions.)
- After parent solutions have mated, what do you do with them? (If they don't die off at some point, they become a drain on the resources in the experimental environment.)
- How long do you let the experiment run?
- How do you determine success?

Micu gave three examples of how evolutionary computing would work in insurance.

Using brainpower to solve problems is, well, a no-brainer. Using evolutionary methods is a bit more revolutionary.

1. **A program that extracts common terms from the text portion of notes taken by underwriters or claims adjusters.**

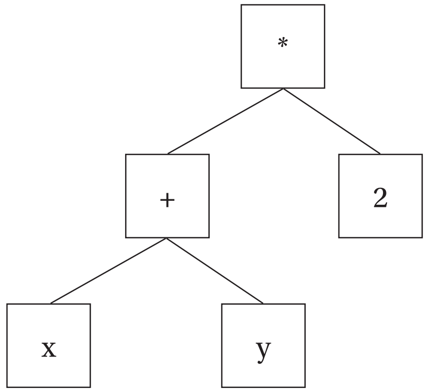
   There are too many combinations to use brute force to determine which phrases are likely to be correlated with a high value for the ultimate loss, but an evolutionary search would allow these phrases to emerge through natural selection.

2. **A program that determines which combination of rating factors gives the best profit, subject to some volume constraint.**

   Gradient-based methods can find local optimums, but "if they get stuck in such a local optimum, there is no way" to reach the overall best result, Micu said. But a type of evolutionary algorithm, termed "differential evolution," can get unstuck from local optimums and explore more of the search space.

3. **A program that creates the tightest fit for a set of data.**

Micu's example used parse trees for arithmetic expressions to perform regression on loss ratios. A parse tree, in this case, is a way to represent an arithmetic expression in a root-and-branch design, resembling the factor trees you learned in elementary school. A parse tree of the expression 2(x + y), for example might be



The parse tree can be thought of as a chromosome, in which each node is analogous to a gene. Crossover occurs when pieces of one parent's parse tree, x + y in this example, swap out with a similar piece from the other parent.

The formula that best explained the observed loss ratio was, in Micu's model, dense and complicated. It looked randomly generated. The same variables appeared more than once.

Evolutionary computing isn't necessarily guaranteed to outperform problem-specific algorithms in all instances, Micu said. But it can expand the toolbox of solutions actuaries use, particularly when problems are complex. Using evolutionary algorithms in conjunction with other machine-learning algorithms can lead to superior models.

James P. Lynch, FCAS, is chief actuary and director of research for the Insurance Information Institute.

Advanced Analytics: Getting Under Way BY JIM LYNCH

Big data and advanced analytics have expanded the actuarial toolkit, at least in theory.

But the theory has been a challenge to put into practice. It can be hard to figure out which data set to tap and which analytical tools to apply and which problems to solve. CAS members heard ideas on how to approach predictive analytics at the recent CAS Ratemaking and Product Management seminar at a session titled "Applying 'Big Data' Analytics in the Insurance Sector."

The insurance industry seems interested in analytics, according to A.M. Best's most recent insurer survey. Nearly 32 percent named advanced analytics as the greatest opportunity before the industry, far more than No. 2 — increasing the use of mobile apps. Applying big data to insurance problems was No. 3.

Clearly, insurers see the value of sophisticated analysis, and as quantitative experts, actuaries would seem likely to deliver the promise of the data into management's hands.

But wielding the new tools is a tricky business. During the session, seminar speaker Jason Rodriguez, data analytics manager at Willis Towers Watson, offered ideas on how to approach a big data project as well as tips on how to manage one.

By "big data," Rodriguez referred to enormous data sets that can present challenges in extracting. An example: text messages, where researchers must sift through and combine information that has not been formatted or standardized, such as great = grate = gr8. In other words, the data is dirty.

Moderator Claudine Modlin and speaker Jason Rodriguez.

The data is so complex, he said, "It really requires us to use more advanced techniques … to extract information."

Often machine learning is used to find useful signals in noisy data. It is an analytic system that can learn as it analyzes and is used to clean up the dirty data. That's a step up from traditional programming, in which the computer consistently responds to the data based on a set of rules. Rodriquez explained that with machine learning, you can train an algorithm to distinguish between cats and dogs based on pictures of each, then let the computer by itself classify a new set of pictures. Machine learning tools are critical in creating order from the relative chaos that big data presents.

Once it is tidied up, big data is ready for predictive analytics, the branch of advanced analytics used to make predictions about unknown events.

A company that wants to undertake a project has a series of decisions to make early on. One of these is to decide which business problem to address. Analytics can help with most insurance functions — underwriting policies, detecting fraud or developing sales leads, for example.

Naturally, actuaries think of the potential to price insurance, but analytics could help set claims reserves or guide claims processes, too. Topic modeling is a text-mining tool that extracts meaning from a block of text. Rodriguez said that a topic modeling tool could scour claims notes looking for words like "hospital" or "emergency" and create new predictors that could be used to estimate claim settlement values or the probability that a claim becomes much more complicated over time, which would let a company quickly assign the trickiest claims to the most seasoned adjusters.

"Advanced analytics would help you improve business decisions based on the data available," Rodriguez said.

The company also has to know what it hopes to accomplish. Rodriguez outlined three typical goals for pursuing advanced analytics opportunities:

- Discovery — exploring your data to learn more about your customers, the goal of which is to extract knowledge, not optimize or minimize some measure. A typical application would be to help product developers learn ways to segment customers.
- Optimization — finding a way to improve a defined business process, like improving a rating plan or increasing claims handling efficiency. Here, Rodriguez said a company has to be sure that implementing the results of the analysis would be feasible — for example, a pricing model would have to be able to fit in a point-of-sale rating engine, or a claims triage model would have to be able to be deployed within the claims workflow and systems.
- Automation — teaching a machine to follow a human process. You "take your business decision and turn it into code," he said, creating decisions "that don't need to be monitored except with minimal human input."

Usually a company has several ideas for a project. To pick the best one, Rodriguez recommended considering the following:

Project viability: Is there management buy-in? How complex is the project, and how well will you be able to assess its success? What resources (data, subject experts) are available?

Clearly, insurers see the value of sophisticated analysis, and as quantitative experts, actuaries would seem likely to deliver the promise of the data into management's hands. But wielding the new tools is a tricky business.

Supporting data: What internal data can you use? If you need to add external data, how well can it link up with your data?

Potential impact: How does the project fit into the overall business strategy? How long will it take before the project yields tangible results?

Implementation issues: Will stakeholders accept the implementation solution? How difficult will it be to implement the solution in existing technology?

The most common problems, Rodriguez said, relate to data and expertise. Sometimes analysts can't get the data they need, particularly if it is administered by a third party. Sometimes internal data isn't accurate enough, or there isn't enough to carry out the analysis. Other times, analysts don't have the right sort of training.

In managing the project, Rodriguez recommended what is known as an agile approach. A traditional project takes place over many months, with a lengthy planning phase followed by a lengthy building phase and a lengthy review phase. An agile approach breaks the project into several small pieces, each of which has its own (shorter) planning, building and review phases.

The advantage of the agile approach is that each component project is delivered sooner, bringing value to the company quickly. This ensures that effort is efficiently translated to valuable results even if the company's needs change during the course of the project.

"You deliver value quickly," Rodriguez said, "then add features [in later phases] to continue to improve."

Rodriguez emphasized the need to choose an analytical approach and technology spend that fit the problem at hand, keeping in mind that mature technologies exist to deal with some increasingly common problems. ●

You've Built the Model. Have You Done It Right? BY JIM LYNCH

Building a predictive model — or any complex actuarial model — is a big job. And it leaves you with a big question: Can the model do the work it was created to do?

Answering that question was the focus of a session titled, prosaically, "How to Pick a Better Model," at the CAS Ratemaking and Product Management Seminar in San Diego in March.

Actuaries started building predictive models about 20 years ago to develop rating plans. Now models are spreading to help underwriters, adjusters and other insurance personnel work more efficiently, and actuaries are at the forefront of building those models.

A model needs to be tested to ensure that it doesn't have fatal characteristics, said Hernan Medina, senior principal data scientist at ISO. Some fundamental flaws of a model could include that it might:

- Be underfit, poorly explaining the data it is modeling.
- Be overfit, explaining the data it is modeling well, but failing to predict accurately when given new data.
- Not perform as well as models already in use.

Medina and Dan Tevet, who is an actuary at Liberty Mutual Insurance, discussed the three key issues that tell whether a model is doing its job well:

1. Lift — can the model distinguish between good risks and bad risks?
2. Goodness of fit — does the model do a good job of explaining the data collected?
3. Stability — is the model likely to stand the test of time?

Medina focused on model lift and Tevet discussed goodness of fit and stability. Hoi Leung, director of predictive analytics at AIG, served as moderator.

"Lift" is sometimes called the economic value of a model, Medina said, because it concerns the model's ability to distinguish between risks.

Models are built based on a set of historical data. The analysts noted that to properly build and test a model, the data should be randomly split, usually into two parts.

One part will be the data, called the "training data," that the actuary uses to create the model. The other part contains the "validation data," the data against which the new data will get its first test run. . A good model will predict about as well on the validation data as it did on the training data. Often data are split into three parts. The training and validation samples may be used to fit and evaluate a few new model alternatives, leaving a "holdout" sample to obtain the final selected model.

Hernan Medina

Dan Tevet

Model Lift

"Lift" is sometimes called the economic value of a model, Medina said, because it concerns the model's ability to distinguish between risks. The classic actuarial example might be the way that many carriers assess the value of insurance credit score. As the score goes up, the expected loss of a risk falls.

The relationship just described is quite similar to one of the tests that Medina cited — the loss ratio test. Loss ratios of actual data are ranked by the loss cost the model predicted. If the model is true, the loss ratio rises as the predicted loss cost rises.

This isn't the most statistically rigorous test, Medina said, because the premium underlying the loss ratios are results of the current rating plan, not the proposed plan, and the current plan's inadequacies could distort the analysis. Still, it is the most straightforward way to display data that most insurance professionals can understand.

Actuaries can analyze results better running the model against the validation data and plotting actual results versus what the model predicted. As the prediction rises, the actual results should rise. This type of chart is called a quantile plot.

One could also use a double quantile plot. The term "double" here means that instead of just plotting one set of predictions against actual results, the analyst plots two sets of predictions. Usually one set of predictions is from the current rating plan and the other set comes from the newly created model. Charting the two models together makes it easier to distinguish at a glance which one performs best.

Goodness of Fit

Any statistician will tell you a model needs to fit the data used to design it. The methods of measuring how close the model fits the data are called goodness-of-fit tests, and understanding how they work can help an analyst improve the model.

Statisticians look at how much the model's prediction differs from the actual data point. This is called the "error term" or the residual.

The most common methods of measuring residuals (squared error, absolute error) aren't appropriate for insurance models, Tevet said. Both work best on normally distributed data — data that accumulates into a bell curve.

Insurance data rarely fits a normal distribution, Tevet said. Using standard goodness-of-fit tests can lead to adopting the wrong model.

Tevet said that it is better to look at a measure known as the deviance of a model. This is similar to using squared error. In fact deviance in a normal distribution is measured by the weighted sum of squared error. Other distributions measure the statistic differently. The deviance of the insurance-friendly Tweedie distribution is

$$2\sum_i w_i \Big(y_i \frac{y_i^{1-p} - \mu_i^{1-p}}{1-p} - \frac{y_i^{2-p} - \mu_i^{2-p}}{2-p},$$

where p is the shape parameter of the distribution.

The deviance of model results can be tweaked into a measure known as the deviance residual. This measure shows the amount by which a model missed its target, but has been adjusted so that all of the deviance residuals, taken together, should form a normal distribution. So each deviance residual misses its target by a random amount.

So a visualization of deviance residuals — a scatterplot of them vs. the predicted variable — should look like a random cloud, with no discernable pattern.

Stability

The final steps Tevet discussed involved determining how robust the model is — making sure it is stable (the parameters that drive the model don't change too quickly) and is not overfit (the model does well on the data it was trained on but not so well on anything else).

"You might sacrifice a little bit of lift for a model that is more stable over time," Tevet said.

Models should also be tested "out of time," meaning testing the model on data gathered from a later time period. That is important in insurance, he said, since the training and testing data, both random subsets of a larger data set, might both contain losses from the same catastrophe or harsh winter.

To protect against overfitting, Tevet suggested a technique known as cross-validation. This is an alternative to creating training and validation datasets.

For example, a modeler could split a data set into five equally sized pieces, known as "folds," take a random sample from each, then fit the model five times. Tevet suggested using cross-validation when the data set being modeled is thinly populated.

Another way to improve the value of the dataset is by creating a new data set from it using a process known as bootstrapping. The new data set is created by randomly selecting data with replacement from the old data set.

"Each random sample can be thought of as an alternative reality," Tevet said.

For example, if you had a bag with 100 marbles, some blue and some red, you could create a virtual marble bag by picking a marble from the actual bag, noting its color then putting it back in the bag it came from. After doing this 100 times, you have a virtual marble bag.

The main advantage of bootstrapping is that the modeler can use the results to create statistical confidence intervals, Tevet said. Then the modeler can better tell if the difference between the model's predictions and reality is from a weakness in the model or just due to chance.

In the end, the speakers agreed, the statistical procedures of picking a model have a sound mathematical basis, but the business knowledge that actuaries add to the process is also crucial.

The methods of measuring how close the model fits the data are called goodness-of-fit tests, and understanding how they work can help an analyst improve the model.

EXPLORATIONS BY DAVE CLARK

Estimation of Inverse Power Parameters via GLM

Richard Sherman's 1984 paper "Extrapolating, Smoothing and Interpolating Development Factors" provided a number of useful ideas for working with development patterns and remains a useful resource. The CAS's 2013 Tail Factor Working Party found that the Sherman curve fit "enjoys fairly broad acceptance both with consulting firms and insurance companies."

In *Variance*, Volume 9, Number 2, Jon Evans has two new papers that extend the ideas of Sherman's original paper and show its continued relevance.

The great value of the original Sherman paper is in identifying a curve form that closely fits the sequence of age-to-age (ATA) factors for long-tailed casualty lines. The most basic form is the "inverse power" curve of the following form:

$$ATA_t = 1+a\cdot t^{-b}.$$

In this form, the t represents the development time such that, for example, ATA_{12} represents the age-to-age factor or link ratio from age 12 months to age 24 months. A modest expansion of this formula allows a shift term, c, to be added to the time index, though we will ignore this for the present discussion:

$$ATA_t = 1+a\cdot(t+c)^{-b}.$$

The parameters of the inverse power curve are most frequently estimated by rearranging the formula into a (log) linear form and then applying ordinary least squares formulas for the intercept and slope:

$$ln(ATA_t\text{-}1)=ln(a)+b\cdot ln(1/t).$$

The attraction of this log-linear form is that simple, closed-form solutions can produce the estimated model parameters. Anyone with a spreadsheet can apply the method with little technical knowledge.

Further, the inverse power curve can easily be compared with alternative fitted curves. Sherman gives several examples, with the exponential decay formula being perhaps most familiar:

$$ATA_t = 1+a\cdot e^{-b\cdot t}.$$

The exponential decay formula can be calculated in a similar log-linear form, so that we quickly have alternative fitted curves to compare to our development data:

$$ln(ATA_t\text{-}1)=ln(a)\text{-}b\cdot t.$$

While the mathematical simplicity of the log-linear form is appealing, it creates difficulties in practice. The difficulties were noted in the discussion of Sherman's paper by Lowe and Mohrman (1985). The first difficulty is that the log-transform $ln(ATA_t\text{-}1)$ requires that every ATA factor used in the fit be strictly greater than 1.000. There can be no "negative development" in the actual data, and even factors that are only slightly greater than 1.000 can cause distortions in the fit.

A second problem is that the log-transformed data is a bit more difficult to interpret or explain to the audience receiving the results of the analysis. An age-to-age factor of 1.010 is easily interpreted as a 1 percent increase in loss dollars, but what does $ln(.01)$=-4.605 represent? How do we interpret the -4.605 for our client or explain why we want a fitted line that closely matches this value?

Both of these difficulties are overcome when we instead approach the parameter estimation using generalized linear models (GLM). We can still use the "inverse power" form that fits the insurance patterns so well, but make use of a better technique for the parameter estimation.

The key idea in GLM is that we include a "link function" $g()$ but apply it in inverse form $g^{-1}()$ to the linear combination of the predictor variable(s). Rather than apply a log-transform to the quantity (ATA_t-1), we use an exponential transform on the linear function.

$$ATA_t\text{-}1 = exp(\beta_0+\beta_1\cdot ln(t)) = \mu_t$$

$$a=exp(\beta_0) \quad b=\text{-}\beta_1$$

Using this "log-link" on the right side of the equation rather than applied to the response variable, we avoid any problem with actual negative development. Expected development must still be positive but the actual values being fit need not be. In short, a log-link GLM can handle negative development in the data where a log-linear regression cannot. The GLM approach is more robust.

With the log-link, the "canonical" variance structure is the quasi-Poisson or over-dispersed Poisson (ODP) model. The ODP model assumes that the variance is proportional to the expected value.

The GLM application follows a Poisson quasi-loglikelihood (QLL). The prefix *quasi* means that we are not explicitly assuming a distribution but rather only assuming that the variance is proportional to the variance of the Poisson distribution.

The reader is referred to the 1974 Wedderburn paper for a more complete description of quasi-likelihoods.

For our application the Poisson QLL is given below:

$$QLL = \sum w_t \cdot [(ATA_t\text{-}1) \cdot ln(\mu_t) \text{-} \mu_t].$$

The function allows weights w_t to be included as part of the fitting procedure. Since we typically use dollar-weighted average ATA factors, the weights are naturally set as the sum of the dollars in the column used in the denominator of the ATA calculation

$$ATA_t = \frac{\sum_{i=1}^{n-t} C_{i,t+1}}{\sum_{i=1}^{n-t} C_{i,t}} \quad w_t \sum_{i=1}^{n-t} C_{i,t}.$$

The QLL can be maximized with the "best" parameters β_0 and β_1 using available software. The **glimmix** procedure in SAS will perform the calculation. The **glm.fit** function in R can also be used but requires a fix to allow negative values (see the code by David Firth in the references). More conveniently, a simple iterative routine can be built into a VB function within an Excel spreadsheet (or even — gasp — using Excel's "Solver").

The estimating equations for finding the best model parameters are easily derived:

$$\sum w_t \cdot (ATA_t\text{-}1) = \sum w_t \cdot (\widehat{ATA}_t\text{-}1)$$

$$\sum w_t \cdot (ATA_t\text{-}1) \cdot ln(t) = \sum w_t \cdot (\widehat{ATA}_t\text{-}1) \cdot ln(t).$$

From these estimating equations, we see that GLM estimation is working with the original dollars from the development triangle, and that the fitted values balance to the actual dollars. There is no difficulty when some actual development is negative and no difficulty in interpreting what is being fit.

The GLM can also be expanded for other transforms of the development time index. Instead of the logarithmic transform that creates the inverse power curve, we can use the time index directly to be equivalent to the exponential decay curve.

If the inverse power curve is too thick-tailed and the exponential decay is too thin-tailed, then other transforms are possible. An intermediate form is to use the square root of the development time.

As with the original Sherman paper, these various transforms of the time index represent variations on the same basic model. Using the log-link GLM form simply gives us a more robust method for estimating parameters for the model.

Inverse Power	Exponential Decay	Square Root
$\mu_t = exp(\beta_0 + \beta_1 \cdot ln(t))$	$\mu_t = exp(\beta_0 + \beta_1 \cdot t)$	$\mu_t = exp(\beta_0 + \beta_1 \cdot \sqrt{t})$

David R. Clark, FCAS, MAAA, works for Munich Reinsurance as part of the actuarial research and modeling team in Princeton, New Jersey.

References

Evans, Jonathan Palmer, "A Continuous Version of Sherman's Inverse Power Curve Model with Simple Cumulative Development Factor Formulas," *Variance* 9:2, 2015, pp. 187-195.

Evans, Jonathan Palmer, "Tail Factor Convergence in Sherman's Inverse Power Curve Loss Development Factor Model," *Variance* 9:2, 2015, pp. 227-233.

Firth, D (2003). An amended quasipoisson function for R. Available at http://bit.ly/2paRUXA.

Lowe, S.P. and D.F. Mohrman, "Discussion of 'Extrapolating, Smoothing and Interpolating Development Factors," *Proceedings of the Casualty Actuarial Society* 72, 1985, pp. 182-189.

McCullagh, P. and J.A. Nelder, *Generalized Linear Models*, Chapman & Hall/CRC, second edition 1989.

Sherman, R.E., "Extrapolating, Smoothing and Interpolating Development Factors," *Proceedings of the Casualty Actuarial Society* 71, 1984, pp. 122-155.

"The Estimation of Loss Development Tail Factors: A Summary Report," *Casualty Actuarial Society Forum*, 2013: Fall, Vol. 1, pp. 1-111.

Wedderburn, R. (1974) "Quasi-Likelihood Functions, Generalized Models and the Gauss-Newton Method," *Biometrika* 61(3), pp. 439-443.

IN MY OPINION BY GROVER EDIE, *AR* EDITOR IN CHIEF

Gram's Sugar Cookies and the Secret Ingredient

My wife's grandmother, "Gram," made excellent sugar cookies. My mother and grandmothers also made sugar cookies that were good, but theirs just weren't at the same level of excellence as Gram's.

When I ran a recent web search I got 2.7 million matches for "sugar cookie recipe." There are entire cookbooks devoted to sugar cookies. You wouldn't think the world would miss another sugar cookie recipe, but I would. That's because none of the rest of them are Gram's recipe.

How do I know that? Because Gram's sugar cookies have a "secret ingredient." I have looked up sugar cookies online and found some recipes with that secret ingredient, but none with the amount she prescribed.

Gram had the recipe in her head, which was a good thing, as she had lost most of her eyesight by the time I met her. Wanting her recipe, my wife, Diane, worked with Gram on a few occasions to bake the magical treats. Gram would put the ingredients into a bowl and Diane would remove them and measure how much there was of each ingredient. They did this a few times to get a good average; thus, the recipe was reverse engineered. Once in the oven, the cookies were "done" when Gram smelled they were done, so oven temperature and baking time also had to be documented.

So what does all this have to do with actuarial matters?

You wouldn't think the world would miss another sugar cookie recipe, but I would. That's because none of the rest of them are Gram's recipe.

The point of this column is not about cookies; it is about passing along to others some of our own secrets — not just cookie and other recipes, but recipes for how to design a spreadsheet or to make links more efficient or even how to present the analysis results in such a way that the recipient will be able to understand and accept. You have secrets

The Secret and the Recipe

Gram's secret ingredient was a pinch of ground nutmeg — not the half teaspoon called for in other recipes, but just a pinch. And it was her small fingers that pinched the spice, not those of someone with large hands, so the amount was rather small — just enough to provide that difference I like so much. She would also put half an apple in the cookie jar to keep the cookies from getting hard and crumbly. We tried it, and it works.

Here's Gram's recipe:

Elsie Herron's ("Gram's") Sugar Cookies

2/3 cup softened butter	2/3 teaspoon baking soda
1 cup sugar	2/3 teaspoon baking powder
2 eggs	Pinch of nutmeg
3 tablespoons sour cream	Scant teaspoon of salt
1 teaspoon vanilla	3 ½ to 4 cups sifted flour, depending on humidity

In a large bowl, cream together sugar and shortening. Add one egg at a time to the creamed mixture, beating well after each egg. Add vanilla and sour cream; stir to combine. In a separate bowl, sift dry ingredients together. Add dry ingredients to wet ingredients, mixing thoroughly.

Separate the dough into two halves.

In warm weather, you may need to refrigerate the dough for up to an hour.

Roll one half of the dough onto a floured surface. Keep some extra flour nearby; you will need more as you work with the dough.

Roll dough thinly, about ¼ inch thickness. Before cutting, sprinkle dough with sugar and lightly roll over with rolling pin, pressing the sugar into the dough.

Cut with a cookie cutter.

Bake at 350 degrees for eight minutes on an ungreased cookie sheet.

to share about how to get things done within your organization or how to best deal with some difficult people. These secrets are not just about the technical aspects of our profession, but the human relations side as well. Think about it — if Gram had not spent the time with Diane to make some cookies, and Diane had not documented the process, I wouldn't be able to enjoy Gram's sugar cookies today and neither would our sons or grandchildren.

What tricks of the trade will disappear when you retire unless you share them?

What have you learned that you can pass on, teaching others rather than leaving it to them to discover themselves?

In my opinion, such thoughts are best pondered while eating a sugar cookie.

IT'S A PUZZLEMENT BY JON EVANS

Infinity Within Infinity Within Infinity?

Here is a question that involves three nested levels of infinity, but first here is an explanation of some terms and concepts. A *partition* of the counting numbers 1, 2, 3,... is a collection of subsets such that each number belongs to exactly one subset. Here are some example partitions:

A = { {1}, {2}, {3},...} B = { {1}, {2, 3}, {4, 5}, {6, 7},...}

C = { {1, 2}, {3,4}, {5,6},...} D = { {1, 3, 5,...}, {2, 4, 6,...} }

Partitions A, B and C each have *infinitely many subsets*, but each subset only contains a finite group of numbers. Partition D contains finitely many subsets, but each subset contains *infinitely many numbers*. Partitions A and C are *disjoint* since no subset is present in both of them, but partitions A and B are not disjoint since they both contain the subset {1}. We can also form a *set of partitions*, such as {A, B, C, D}.

Is it possible to construct a set of partitions such that all of the following occur?

- There are infinitely many partitions in the set.
- Any two of the partitions are disjoint.
- Each of the partitions contains infinitely many subsets.
- Every subset of every partition contains infinitely many numbers.

Can you either construct such a set of partitions or prove that such a set is impossible to construct?

Malware Versus Anti-Malware

A computer virus is programmed to make three identical copies of itself and then delete itself. Network anti-malware software has a probability P of destroying any given copy of the virus before it can make the three copies. A single copy is introduced into the network. What is the minimum value of P so that there is 99 percent chance that the virus will be completely eradicated eventually? What is the minimum value of P for a 100 percent chance of eventual complete eradication?

Let Q(P) be the probability that eventually all descendants of a given copy of the virus will be eradicated. Then $Q(P) = P + (1-P)\,[Q(P)]^3$ or the sum of the probability that the anti-malware destroys the virus before it can reproduce and the probability that the virus survives to reproduce but all of the descendants of its three children are eventually destroyed. Solving for P(Q), gives $P(Q) = (Q + Q^2)/(1 + Q + Q^2)$. To get Q = 100% requires that P = 2/3 and to get Q = 99% requires that P = 1.9701/2.9701 = 66.3311%.

Solutions were also submitted by Patrick Allen, Xunchi Chen, Bob Conger and Brad Rosin.

Know the answer? Send your solution to ar@casact.org.

Expertise. Insight. Solutions.

Casualty Actuarial Society
4350 North Fairfax Drive, Suite 250
Arlington, Virginia 22203 USA
Phone: 703-276-3100, Fax: 703-276-3108
www.casact.org

PRESORTED
STANDARD MAIL
U.S. POSTAGE PAID
LUTHERVILLE, MD
PERMIT NO. 171