The 56th Actuarial Research Conference

Parallel Session Abstracts

August 19, 2021

3:15 PM – 4:30 PM

Parallel Session I – Emerging Risk/Insurance Technology

3:15 PM – 3:40 PM
Title: Framework for loss distribution of the operational risk of vehicle-to-vehicle cooperation to marshal traffic via monetization of highway space
Presenter: Petar Jevtic, Arizona State University
Abstract: In this work, we develop a framework for loss distribution of the operational risk of vehicle- to-vehicle cooperation to marshal traffic via monetization of highway space. As a motivating technological solution, we consider the Cooperatively Managed Merge and Pass (CMMP) system proposed by Ford Motor Company. In the USA alone, traffic congestion in terms of time lost imposes substantial opportunity costs, afflicts commuters with stress, and contributes to air pollution. Acting in concert, a vehicle-to-vehicle communications infrastructure, blockchain, smart contracts, and digital assets technologies allow for novel solutions to this long-standing problem. So far, leading automobile manufacturers have offered several technological solutions for real-time monetization of highway space. Unfortunately, often the technology stack required has cyber vulnerabilities leaving these solutions open to misuse. In that context, understanding the operational risk via its loss distribution is of high importance for operators and insurers. That is why we pioneer the approach that contextualizes the problem in the probabilistic graph- theoretical framework by using percolation models. In the case of adverse materialization of cyber risk and its spread through a vehicular network, we allow for spatially heterogeneous loss topology. In the context of the USA, we provide instructive numerical examples.

3:40 PM - 4:05 PM
Title: Cost-sensitive Multi-class AdaBoost for Understanding Driving Behavior with Telematics
Presenter: Emiliano A. Valdez, University of Connecticut
Abstract: Powered with telematics technology, insurers can now capture a wide range of data, such as distance traveled, how drivers brake, accelerate or make turns, and travel frequency each day of the week, to better decode driver's behavior. Such additional information helps insurers improve risk assessments for usage-based insurance (UBI), an increasingly popular industry innovation. In this article, we explore how to integrate telematics information to better predict claims frequency. For motor insurance during a policy year, we typically observe a large

proportion of drivers with zero claims, a less proportion with exactly one claim, and far lesser with two or more claims. We introduce the use of a cost-sensitive multi-class adaptive boosting (AdaBoost) algorithm, which we call SAMME.C2, to handle such imbalances. To calibrate SAMME.C2 algorithm, we use empirical data collected from a telematics program in Canada and we find improved assessment of driving behavior with telematics relative to traditional risk variables. We demonstrate the competitiveness of our algorithm against other models that can handle class imbalances. The sampled data on telematics were observations during 2013-2016 for which 50,301 are used for training and another 21,574 for testing. Broadly speaking, the additional information derived from vehicle telematics helps refine risk classification of drivers of UBI. *This is joint work with Banghee So and Jean-Philippe Boucher.*

4:05 PM – 4:30 PM
Title: Extending the Insurance Forecasting Time Horizons from Years to Decades
Presenter: Stephen Kolk, Kolkulations LLC
Abstract: The frequency and intensity of major perils including Hurricanes, Floods, Tornadoes, and Wildfires is increasing slowly. Global warming is causing the extreme catastrophes risk stirred up by these perils to intensify. Traditional insurance ratemaking, which makes forecasts just a few years in the future, doesn't keep pace with climate changes that multiply and worsen all types of catastrophes. This paper will present a simple effective tool, the LOESS method, to find the trend signal in noisy catastrophe data. And it will discuss how it is necessary and essential to extend Property/Casualty forecasting time horizons from a few years to decades to bring accelerating risks into focus and keep pace with the extreme events' growth.

Parallel Session II – Actuarial Education

3:15 PM – 3:40 PM
Title: Society of Actuaries Education Update
Presenter: Stuart Klugman, Society of Actuaries
Abstract: The Society of Actuaries has made significant changes in its education requirements as part of its Long Term Growth Strategy. At this session steps taken to date will be reviewed.

3:40 PM – 4:05 PM
Title: CAS Admissions and Education initiatives
Presenter: Ken Williams, Casualty Actuarial Society
Abstract: The Covid-19 pandemic led to many changes in the last 12 months in the CAS credentialing process. This session will provide an update for the changes in the last 12 months, as well as what is planned for the future.

4:05 PM – 4:30 PM
Title: Transforming CIA Education

Presenter: Bruce Jones, Canadian Institute of Actuaries
Abstract: The nature of work and the workforce has been transformed with the global pandemic. The CIA has leveraged this disruption in many ways over the past 12 months, and one of the most significant has been in our education system. The way we educate and qualify actuaries in Canada has fundamentally and permanently changed. As a result, we are excited to share details regarding the development of a new education system that will make the Canadian actuarial profession more flexible, attractive, and accessible through three qualification pathways designed to accommodate the diversity of candidates for membership in the CIA.

Parallel Session III – Data Science and Applications

3:15 PM-3:40 PM
Title: Self-Organizing Maps to Examine Determinants of Loneliness in Older Adults
Presenter: Seulkee Yun, University of Wisconsin-Madison
Abstract: Loneliness can be defined as feeling sad about being alone and isolated (www.vocabulary.com/dictionary). Feelings of loneliness have been shown to have an adverse effect on one's well-being, physical and mental health and mortality. In this work, we use a machine learning method called Self-Organizing Maps (SOM) as an alternative tool to identify the possible characteristics that could be used to determine subgroups suffering severe loneliness. Classic statistical methods have been used to evaluate determinants of loneliness. However, they usually cannot access the interrelationship of the variables that are associated with different levels of loneliness. SOMs are a neural network that uses unsupervised competitive learning that performs a powerful visual friendly analysis to identify patterns in the data structure. SOM maps nonlinear statistical relationships between high-dimensional data into low dimensional grids while maintaining its original spatial relations in data observations. In this study, we analyze 2016 Health Retirement Study (HRS) data. The HRS dataset is unique because it includes information that are usually unavailable from other national survey data. First, HRS includes not only direct questions whether respondent is lonely, but also questions that indirectly quantify loneliness. Second, HRS contains information on the elderly population living in nursing home, which is important subgroup to compare to the community-dwelling elderly population.

3:40 PM – 4:05 PM
Title: Pricing and Reserving Wildfire Insurance in North America
Presenter: Roba Bairakdar, Concordia University
Abstract: The wildfires in Fort McMurray (Alberta, Canada, May 2016) and the recent wildfires in California (2017-2020) burnt millions of acres, generated billions of dollars in insured losses and alerted the public and insurers to this risk. Homeowners' insurance policies in fire prone areas can be a very risky business that some insurers may not be willing to undertake, resulting

in a transfer of this business to expensive specialty markets. Wildfire risk features strong spatial dependence and heterogeneity. Spatial heterogeneity implies that ratemaking territories should be consistent with local wildfire risk whereas spatial dependence implies that reserves and risk loading on premiums should account for lack of perfect diversification. Random forest regression is used to create an actuarial spatial model for the likelihood of wildfire occurrence over a fine grid map of North America to help assess the loss distribution at the individual and portfolio levels. The model considers natural sparks such as a lightning strike or human activity as triggers for a wildfire, in addition to climate variables and land use as possible pre-existing conditions. The model provides actuaries with the capabilities to price, reserve and/or manage the financial risk from wildfires.

4:05 PM – 4:30 PM
Title: Ultimate Loss Development using Bidirectional LSTMs
Presenter: Lahiru Somaratne, University of Nebraska – Lincoln
Abstract: Loss development has always been at the core of Property and Casualty Insurance. Numerous traditional methods have been developed over the years to tackle the issue of how a loss faced by an insurance company will mature and what will be the ultimate loss to the said insurance firm. Recognizing that loss development can be viewed as a numerical sequence prediction problem, this presentation explores how loss development can be modeled using deep learning. A variation of RNN (using Bi-directional LSTM cells), is used to predict ultimate losses. Deep learning requires comparatively more data than some of the more conventional machine learning methods, therefore the losses of multiple companies are predicted in tandem, to boost individual predicted loss accuracy. The results of this method are compared with that of the 'chainladder' Python package, an already existing broader loss reserving implementation.

6:00 PM - 7:15 PM

Parallel Session I – Finance

6:00 PM – 6:25 PM
Title: On Complex Economic Scenario Generators: Is Less More?
Presenter: Jean Francois Begin, Simon Fraser University
Abstract: This research proposes a complex economic scenario generator that nests versions of well-known actuarial frameworks. The generator estimation relies on the Bayesian paradigm and accounts for both model and parameter uncertainty via Markov chain Monte Carlo methods. So, to the question is less more? We answer maybe, but it depends on your criteria. From an in-sample fit perspective, on the one hand, a complex economic scenario generator seems better. From the conservatism, forecasting, and coverage perspectives, on the other hand, the situation is less clear: having more complex models for the short rate, term structure,

and stock index returns is clearly beneficial. However, that is not the case for inflation and the dividend yield.

6:25 PM – 6:50 PM
Title: The Reinforcement Learning Kelly Strategy
Presenter: Ruihong Jiang, University of Waterloo
Abstract: The full Kelly portfolio strategy's deficiency in the face of estimation errors in practice can be mitigated by fractional or shrinkage Kelly strategies. This paper provides an alternative, the RL Kelly strategy, based on a reinforcement learning (RL) framework. RL algorithms are developed for the practical implementation of the RL Kelly strategy. Extensive simulation studies are conducted, and the results confirm the superior performance of the RL Kelly strategies.

6:50 PM – 7:15 PM
Title: Risk-Constrained Portfolio Choice via Quantiles
Presenter: Michael Zhu, University of Waterloo
Abstract: In problems of budget-constrained portfolio choice in complete markets with law-invariant preferences and linear pricing, it is well-known that the optimal investment leads to a wealth profile that is anti-comonotonic with the state-price density. This result provides the foundation of the so-called quantile formulation approach to portfolio choice problems. In this paper, we contribute to the literature on non-convex portfolio choice problems with non-EU preferences, by extending several models of optimal asset allocation under non-EU preferences to incorporate non-linear pricing and non-linear risk measurement constraints. Specifically, we consider Rank-Dependent Utility models, Yaari's Dual Theory models, as well as other behavioral decision-making models; as well as a budget constraint based on non-linear pricing, and a non-linear risk measurement constraint. We employ a quantile formulation approach and a relaxation approach to provide a closed-form characterization of optimal investment strategies. We illustrate our results with a few numerical examples.

Parallel Session II – Retirement/Pension

6:00 PM – 6:25 PM
Title: It's RILA Time: An Introduction to Registered Index-Linked Annuities
Presenter: Thorsten Moenig, Temple University
Abstract: Registered index-linked annuities (RILAs) are increasingly popular equity-based retirement savings products offered by U.S. life insurance companies. They combine features of fixed-index annuities and traditional variable annuities (TVAs), offering investors equity exposure with downside protection in a tax-deferred setting. This article introduces RILAs to the academic literature by describing the products' key features, developing a general pricing

model, and deriving the providers' hedging strategy by decomposing their liabilities into short-term European options.

Numerical illustrations show that RILAs offer investors similar risk profiles (in the long run) as TVAs with maturity guarantees, and that many products currently sold appear to be priced quite favorably for investors. For providers, RILAs may be a preferable alternative or complement to TVAs as they greatly simplify the management of the embedded equity risk and can naturally reduce the TVA capital requirements. These features position RILAs as a viable long-term solution for this product space.

6:25 PM – 6:50 PM
Title: Valuation of guaranteed minimum maturity benefits under generalized regime-switching models using the Fourier Cosine method
Presenter: Yang Shen, University of New South Wales (UNSW Sydney)
Abstract: This paper presents a flexible valuation approach for variable annuity (VA) contracts embedded with guaranteed minimum maturity benefit (GMMB) riders written on an underlying fund that evolves according to a general regime-switching framework. Unlike the traditional regime-switching analysis, only allowing model parameters to change upon regime switches, our approach allows, more importantly, model structures to vary. With only mild assumptions on the characteristic function of the log-stock price, our framework enables the study of fundamental features of the market movement, such as stochastic volatility, interest rates and jumps, on the price of GMMB. This novelty idea is illustrated by three-regime models whose environments can be characterized by either the geometric Brownian motion process, double exponential process or Heston's stochastic volatility process. Two versions of the GMMB riders are considered, including a fixed or roll-up guarantee and a ratchet geometric average guarantee. With the Fourier Cosine (COS) method that utilizes characteristic functions, explicit valuation expressions for various contracts are derived, and numerical illustrations performed analyzing the efficiency of the approach in terms of computational speed and accuracy. The paper makes a unique contribution by presenting regime-dependent bounds and a mathematical framework for determining the optimal grid points required for the COS method to achieve a specific level of accuracy. Numerical experiments for the valuation framework reveal that as the likelihood of regime shifts increases, the price difference of VA contracts with different initial regimes diminishes, which is consistent with financial intuition.

6:50 PM – 7:15 PM
Title: Target Volatility Strategies for Group Self-Annuity Portfolios
Presenter: Jonathan Ziveyi, University of New South Wales (UNSW Sydney)
Abstract: While the current pandemic is causing mortality shocks the world over, the management of longevity risk remains a major challenge for both individuals and institutions. It is high time there be private market solutions designed for efficient longevity risk transfer

among various stakeholders such as individuals, pension funds, annuity providers among others. From individuals' point of view, appealing features of post-retirement solutions include stable and satisfactory benefit levels, flexibility, meeting bequest preferences and low fees. Group Self-Annuitization (GSA) arrangements, popular in some markets, offer similar features, at least to some extent. However, GSA participants retain all risks, as a pool. While idiosyncratic longevity risk can be offset based on pooling arguments, the benefits which participants will receive from the GSA are exposed to financial and aggregate longevity risk. Financial risk requires appropriate investment strategies. This paper devises target volatility strategies for GSA portfolios. We compare benefit profiles emerging under static and dynamic underlying fund investment strategies. The pool funds are invested in a combination of deterministic cash account and equity whose dynamics evolve according to the Heston stochastic volatility process. Benefit profiles are assessed by analyzing various quantiles and alternative strategies involving varying equity compositions are presented. The case of death benefits is included, and the fund dynamics analyzed by assessing resulting investment returns and the mortality credits. Overall, higher living benefit profiles are obtained under a dynamic target volatility strategy. From the analysis performed, a trade-off between the equity proportion and the impact on the lower quantile of the living benefit amount emerges, which suggests an optimal proportion of equity composition.

Parallel Session III – Risk Measures/Cyber Risk

6:00 PM – 6:25 PM
Title: Smoothed Quantiles for Measuring Discrete Risks
Presenter: Vytaras Brazauskas, University of Wisconsin-Milwaukee
Abstract: Many risk measures can be defined through the quantile function of the underlying loss variable (e.g., a class of distortion risk measures). When the loss variable is discrete or mixed, however, the definition of risk measures has to be broadened, which makes statistical inference trickier. To facilitate a straightforward transition from the risk measurement literature of continuous loss variables to that of discrete, in this paper we study smoothing of quantiles for discrete variables. Smoothed quantiles are defined using the theory of fractional or imaginary order statistics, which was originated by Stigler (1977). To prove consistency and asymptotic normality of sample estimators of smoothed quantiles, we utilize the results of Wang and Hutson (2011) and generalize them to vectors of smoothed quantiles. Further, we thoroughly investigate extensions of this methodology to discrete populations with infinite support (e.g., Poisson and zero-inflated Poisson distributions). Finally, applications of smoothed quantiles to risk measurement (e.g., estimation of distortion risk measures such as value-at-risk, conditional tail expectation, and proportional hazards transform) are discussed and illustrated using automobile accident data. Comparisons between the classical (and linearly interpolated) quantiles and smoothed quantiles are performed as well.

6:25 PM – 6:50 PM
Title: Some results on the risk capital allocation rule induced by the Conditional Tail Expectation risk measure
Presenter: Nawaf Mohammed, York University
Abstract: Risk capital allocations (RCAs) are an important tool in quantitative risk management, where they are utilized to, e.g., gauge the profitability of distinct business units, determine the price of a new product, and conduct the marginal economic capital analysis. Nevertheless, the notion of RCA has been living in the shadow of another, closely related notion, of risk measure (RM) in the sense that the latter notion often shapes the fashion in which the former notion is implemented. In fact, as the majority of the RCAs known nowadays are induced by RMs, the popularity of the two are apparently very much correlated. As a result, it is the RCA that is induced by the Conditional Tail Expectation (CTE) RM that has arguably prevailed in scholarly literature and applications.

Admittedly, the CTE RM is a sound mathematical object and an important regulatory RM, but its appropriateness is controversial in, e.g., profitability analysis and pricing. In this paper, we address the question as to whether or not the RCA induced by the CTE RM may concur with alternatives that arise from the context of profit maximization. More specifically, we provide exhaustive description of all those probabilistic model settings, in which the mathematical and regulatory CTE RM may also reflect the risk perception of a profit-maximizing insurer.

6:50 PM – 7:15 PM
Title: Cyber Risk Assessment for Capital Management
Presenter: Linfeng Zhang, University of Illinois at Urbana-Champaign
Abstract: Cyber risk is an omnipresent risk in the increasingly digitized world that is known to be difficult to quantify and assess. Despite the fact that cyber risk shows distinct characteristics from conventional risks, most existing models in the insurance literature have been based on frequency-severity analysis, which was developed for classic property and casualty risks. In contrast, the cybersecurity engineering literature employs different approaches under which cyber incidents are viewed as threats or hacker attacks acting on a particular set of vulnerabilities. There appears a gap in cyber risk modeling between engineering and actuarial science literature. This paper presents a novel vulnerability-threat model to capture this unique dynamic of cyber risk and to predict loss distributions given a particular cybersecurity profile. The analysis leads to a new tool for allocating a company's resources between cybersecurity investment and loss-prevention reserves.

August 20, 2021

3:15 PM – 4:30 PM

The 56th Actuarial Research Conference

Parallel Session I – Life Insurance

3:15 PM – 3:40 PM
Title: Stochastic mortality dynamics driven by mixed fractional Brownian motion
Presenter: Kenneth Zhou, Arizona State University
Abstract: Recently, the long-range dependence (LRD) of mortality dynamics has been identified and studied in the actuarial literature. The non-Markovian feature caused by LRD can raise new research questions in actuarial pricing and risk management. This paper proposes a new modeling approach that uses a combination of independent Brownian motion and fractional Brownian motion to achieve a more flexible setting on LRD for mortality dynamics. The closed-form solutions of survival probabilities are derived for valuation and hedging purposes. To obtain mortality sensitivity measures in the presence of LRD, we develop a novel derivation method using directional derivatives. Our method is flexible in the sense that it can not only reflect the effect of LRD on mortality sensitivities, but also include some existing sensitivity measures as special cases. Finally, we provide empirical illustrations to compare the performance of different sensitivity measures in natural hedging of longevity risk.

3:40 PM – 4:05 PM
Title: Metamodeling for Variable Annuity Valuation: What works and what does not
Presenter: Xiaochen Jing, University of Wisconsin – Madison
Abstract: Variable Annuities have become popular retirement products with various options of guarantees, but their complex design also makes liability management a difficult task for insurers. There have been several dozen papers published in the past years on exploring the use of statistical learning and metamodeling approaches for Variable Annuity valuation and risk management in the actuarial science and quantitative finance literatures. However, they all focus on specific techniques in the context of synthetic data. In this paper, I investigate the effectiveness of metamodeling approaches with different experimental designs and metamodels with real-world Variable Annuity contracts. In particular, I use textual analysis to extract and formulate value-related features and develop a flexible and comprehensive simulation-based scheme for Variable Annuity valuation. I find that (1) real-world variable annuity contracts are very complex and the intricate relations between their valuation and features are difficult to obtain. And (2) the overall performance of a metamodeling method depends on the employed machine learning methods as well as the sample size---though not substantially on the sampling methods. Both improve performance at the cost of longer runtime.

4:05 PM – 4:30 PM
Title: Common Mortality Trend Model and Mortality Prediction
Presenter: Yechao Meng, University of Waterloo

Abstract: In the paper, we propose a framework to allow populations at disparate mortality development stages to be contemporaneously taken into account through a bivariate-population mortality system with an underlying process of common mortality trend (CMT). As one of the main contributions of the paper, the CMT model is used as a base learner in a computationally friendly framework to effectively borrow information from multiple populations for the mortality prediction of a target population through a well-established procedure of model averaging. Empirical studies with the Human Mortality Database (HMD) confirm that the proposed CMT-based prediction framework yields a substantial improvement in prediction performance over a set of benchmark prediction models. Furthermore, these empirical studies reveal that a key parameter introduced in the CMT model is able to characterize the development stage of one population relative to the other in the CMT model.

Parallel Session II – Actuarial Education

3:15 PM- 3:40 PM
Title: A comparison of Macaulay approximations
Presenter: Stefanos Orfanos, Georgia State University
Abstract: We examine a number of approximation formulas for the net present value of common cash flow streams —such as annuities with or without negative payments and bonds with or without embedded options— that utilize their Macaulay duration and convexity. Specifically for the case of bonds with embedded options, we consider the notions of effective duration and convexity instead. The emphasis is in assessing the accuracy of the various approximation formulas under different scenarios. To simplify the exposition, we only consider parallel shifts to flat yield curves.

3:40 PM – 4:05 PM
Title: Loss Data Analytics: Short Course
Presenter: Edward Frees, University of Wisconsin-Madison, Australian National University
Abstract: One of the silver linings of the COVID pandemic is that it has forced actuarial academics to move beyond traditional face-to-face teaching methods, sometimes known as the "sage on the stage," and consider supplement teaching with computer-driven tools. In this talk, we describe an open source project being developed by 15 collaborators from 4 countries. The purpose of our pilot project, under the auspices of ASTIN, the non-life section of the International Actuarial Association, is to provide a series of tutorials on Loss Data Analytics. The development site is https://openacttextdev.github.io/LDACourse1/. We hope that attendees of the talk will learn from this resource, encourage students and colleagues to also do so, and consider contributing to our efforts as we seek to provide open educational resources to the actuarial profession.

4:05 PM – 4:30 PM

Title: Teaching and Defining Problem Difficulty For Actuarial Exams
Presenter: Russell Hendel, Towson University
Abstract: This presentation proposes a solution to the student actuarial paradox, "I mastered the syllabus but failed the exam," by utilizing the two-dimensional syllabus vehicle which pedagogically focuses on the two dimensions of content and problem difficulty. A contribution of this presentation is the PIR characterization of three levels of problems difficulty: Plug-ins (easy), Integrated (intermediate), and Research-type(hard) questions. PIR was developed by examining difficulty levels in SOA exams, and problems of several software vendors. The author's students routinely score high on the SOA FM exam; it is hoped that these methods will assist other instructors in their quest for good pedagogy.

Very briefly: "Plug-ins," refer to multi-part problems using the same syllabus module, for example, calculating the price of an annuity that is first increasing and then level; "Integrated problems" specifically refer to multi-part problems where the two or more problem parts do not belong to the same syllabus component and are not expected to be "seen" together, for example, calculating an extrema (calculus module) in the difference between a level and inflation adjusted perpetuity; "Research-type questions" refer to multi-part problems that integrate multiple syllabus components in a way never seen in classroom illustrative examples.

PIR can be applied to: i) Pedagogy: Inform students that besides content they are being taught organizational principles, ii) Evaluation: Examinations should reflect certain percentages of each PIR type and iii) Remediation: direct requested student help to formula memorization (for plug-ins), but to organizational techniques for integrated problems, and to problem breakup for research problems.

4:30 PM – 4:55 PM
Title: All That You Can't Leave Behind: Lessons from Pandemic Teaching
Presenter: Diana Skrzydlo, University of Waterloo
Abstract: Teaching online during the pandemic was different in a lot of ways – some bad and some good. We needed to rethink our priorities, the way we delivered content, and the way we interacted with and assessed our students. Now we have the opportunity to leverage all that work into a superior educational experience when we return to the classroom. In this session I'll share some ideas about the things we shouldn't leave behind and how to adapt them going forward. This session can be considered a follow-up to the August 5 SOA Town Hall discussion "Bringing the Best of Pandemic Teaching Back to Class" but will also be completely accessible to people who did not attend it.

Parallel Session III – COVID-19

3:15 PM – 3:40 PM

Title: Actuarial analysis on a deterministic SEIR model
Presenter: Hee Seok Nam, Kettering University
Abstract: In this presentation, we consider a classical epidemiological compartment model from actuarial view points. Focusing on a deterministic SEIR model, infectious disease insurance plans are investigated analytically to obtain properties and relations between various quantities such as annuities, lump sum benefit insurances and level net premiums.

3:40 PM – 4:05 PM
Title: Epidemiological compartment models and premium calculation
Presenter: Vajira Manathunga, Middle Tennessee State University
Abstract: With COVID, epidemiologic compartmental models such as SIR became very popular in the mathematical community. However, there are very few research articles that try to connect these compartmental models to actuarial applications. In this talk, we consider the following insurance/annuity products.
Plan 1: Premium paid while in susceptible state and benefit collected in annuity form while insured in an infected state
Plan 2: Premium paid while in susceptible state and benefit collected as a lump sum when insured enter into an infected state
Plan 3: Annuity or lump sum payment would be triggered when the number of infected is more than the pre-determined number.
Plan 1 and Plan 2 were discussed in [1] under the SIR deterministic model but not under any other compartment models or stochastic models. Plan 3 considers a situation where payments would be triggered if the number of infected people exceeds some pre-determined number. This is similar to a business interruption plan for a virus outbreak. In this talk, we plan to discuss the premium calculation of plan one, plan two, and plan three under several epidemiological compartment models.
[1] Feng, R., & Garrido, J. (2011). Actuarial applications of epidemiological models. North American Actuarial Journal, 15(1), 112-136.

4:05 PM – 4:30 PM
Title: Healthcare loss triangle using epidemiology model
Presenter: Longhao Jin, University of Illinois Urbana-Champaign
Abstract: The impact of the COVID–19 pandemic in the health insurance industry have primarily been felt through increase in health spending and uncertainty in projecting future medical claims. One significant consideration is that health insurers pay different benefits to the infected individuals over their course of infection. However, traditional actuarial models lack the flexibility and robustness to describe the rapidly changing environment during a pandemic. To that end, we explore the epidemiology literature and combine some of the commonly used models with actuarial methodologies. We adopt the infection-age epidemic model and design epidemic insurance plans based on this model. Compared to classical life insurance, healthcare

insurer's reserve function exhibit quite different patterns over the course of a pandemic and so we formulate conditions under which various reserve shapes arise. Inspired by well-known loss reserving techniques, we introduce a loss triangle, derived from the infection-age model, which is capable of predicting future unpaid healthcare liabilities during a pandemic.

6:00 PM – 7:40 PM

Parallel Session I – Risk Measures/Risk Management

6:00 PM – 6:25 PM
Title: Ordered Risk Aggregation under Dependence Uncertainty
Presenter: Yuyu Chen, University of Waterloo
Abstract: We study the aggregation of two risks when the marginal distributions are known and the dependence structure is unknown, under the additional constraint that one risk is no larger than the other. Risk aggregation problems with the order constraint are closely related to the recently introduced notion of the directional lower (DL) coupling. The largest aggregate risk in concave order (thus, the smallest aggregate risk in convex order) is attained by the DL coupling. These results are further generalized to calculate the best-case and worst-case values of tail risk measures. In particular, we obtain analytical formulas for bounds on Value-at-Risk. Our numerical results suggest that the new bounds on risk measures with the extra order constraint can greatly improve those with full dependence uncertainty.

6:25 PM – 6:50 PM
Title: New Class of Distortion Risk Measures and Their Estimation
Presenter: Xiwen Wang, Central Michigan University
Abstract: In this paper, we present a new method to construct new classes of distortion functions. A distortion function maps the unit interval to the unit interval and has characteristics of a cumulative distribution function. The method is based on the transformation of an existing non-negative random variable whose distribution function, named the generating distribution, may contain more than one parameter. The coherency of the resulting risk measure is ensured by restricting the parameter space on which the distortion function is concave. We study the cases when the generating distributions are exponentiated exponential and Gompertz distributions. Closed-form expressions for risk measures are derived for uniform, exponential, and Lomax losses. Numerical and graphical results are presented to examine the effects of parameter values on the risk measures. We then propose a simple estimator of risk measures and conduct simulation studies to compare and demonstrate the performance of the proposed estimator for various losses.

6:50 PM - 7:15 PM
Title: On the Modeling and Pricing of Storm CAT Bond

Presenter: Shimeng Huang, University of Wisconsin-Madison
Abstract: Both the frequency and intensity of weather-related catastrophes such as storms and floods, have been increasing due to climate change such as global warming. This leads to rising storm catastrophe risks faced by the property-casualty insurance industry. In this paper we propose an index-based storm catastrophe (CAT) bond for reinsurers to hedge risks related to storm losses. We discuss the modeling and pricing of the proposed storm CAT bond using historical storm losses at the county level in Florida. We model the losses using a Gamma- two-part autoregressive (2PAR) distribution as the marginal model, and the spatio-temporal vine copula as the dependence model to capture the characteristics of storm losses. We price the proposed security using actuarial models. Our empirical results show the proposed bond can stabilize the reinsurer's cash flows by significantly reducing the variance of the reinsurer's wealth and create attractive returns to investors by offering attractive coupon rates.

7:15 PM - 7:40 PM
Title: Optimizing distortion riskmetrics with distributional uncertainty
Presenter: Qiuqi Wang, University of Waterloo
Abstract: Optimization of distortion riskmetrics with distributional uncertainty has wide applications in finance and operations research. Distortion riskmetrics include many commonly applied risk measures and deviation measures, which are not necessarily monotone or convex. One of our central findings is a unifying result that allows us to convert an optimization of a non-convex distortion riskmetric with distributional uncertainty to a convex one, leading to great tractability. The key to the unifying equivalence result is the novel notion of closedness under concentration of sets of distributions. Our results include many special cases that are well studied in the optimization literature, including but not limited to optimizing probabilities, Value-at-Risk, Expected Shortfall, and Yaari's dual utility under various forms of distributional uncertainty. We illustrate our theoretical results via applications to portfolio optimization, optimization under moment constraints, and preference robust optimization.

Parallel Session II – Credibility Theory/Credibility Model

6:00 PM – 6:25 PM
Title: Limited-Fluctuation Credibility under Uncertain Priors
Presenter: Michael Baron, American University
Abstract: The classical notion of credibility is revisited, with an additional model explaining uncertainty in the prior distribution. This reflects practical situations where the prior distribution of parameters in the actuarial frequency-severity model is not fully deterministic, for example, when it is estimated from past experience. Assumption of a fully credible prior experience has been pointed as a shortcoming in recent actuarial literature.

When both components of the compromise estimator are subject to uncertainty, the limited-fluctuation credibility approach can result in three outcomes - there may be full credibility, partial credibility, or no credibility at all.

Three methods of evaluating credibility factors are proposed and compared, based on three different interpretations of limited fluctuation under the new general setting. Special methods are elaborated for dealing with heterogeneous risk groups.

This research was supported by The Actuarial Foundation.

6:25 PM – 6:50 PM
Title: Hawkes Processes in Dynamic Credibility
Presenter: Bin Zou, University of Connecticut
Abstract: In the standard credibility theory, claim frequency is modeled by a series of independent discrete distributions; while empirical insurance data often exhibit strong dependency on the past claim history. To address this practical issue, we propose to apply a general Hawkes process to model the claim frequency in a dynamic credibility framework. The results show that such a Hawkes-driven credibility framework outperforms the benchmark static credibility model.

6:50 PM – 7:15 PM
Title: Approximation of zero-inflated Poisson credibility premium via variational Bayes
Presenter: Himchan Jeong, Simon Fraser University
Abstract: We explore a way to approximate credibility premium for claims frequency that follows zero-inflated Poisson distribution via variational Bayes. The proposed method enables insurance companies to capture both zero-inflation and unobserved heterogeneity of policyholders simultaneously with modest computation costs. A simulation study and empirical analysis support the novelty of the proposed method based on prediction performances and computation time compared to the benchmarks.

7:15 PM – 7:40 PM
Title: Bayesian credibility under a bivariate conjugate prior on frequency and severity
Presenter: Eric Cheung, University of New South Wales (UNSW Sydney)
Abstract: In this paper, we propose a credibility model in which the (unobservable) risk profiles of the claim frequency and the claim severity are dependent. Given the risk profiles, the (conditional) marginal distributions of frequency and severity are assumed to belong to the exponential family. A bivariate conjugate prior is proposed for the risk profiles, where the dependency is incorporated via a factorization structure of the joint density. The bivariate posterior is derived, and in turn the Bayesian premium for the aggregate claim is given along with some results on the predictive joint and marginal distributions involving the claim number

and the aggregate claim in the next period. The model covers various useful bivariate conjugate priors in relation to mixed Erlang, gamma mixture, Farlie-Gumbel-Morgenstern (FGM) copula, and bivariate beta. A numerical example will be provided to illustrate the effect of dependence on the results.

Parallel Session III – Data Science and Applications

6:00 PM – 6:25 PM
Title: Imbalanced Learning using Actuarial Modified Loss Function in Tree-Based Models
Presenter: Changyue Hu, University of Illinois at Urbana Champaign
Abstract: Tree-based models have gained momentum in insurance claim loss modeling; however, the point mass at zero and the heavy tail of insurance loss distribution pose the challenge to apply traditional methods directly to claim loss modeling. With a simple illustrative dataset, we first demonstrate how the traditional tree-based algorithm's splitting function fails to cope with data with a large proportion of zeros. To address the imbalance issue presented in loss modeling, this paper aims to modify the traditional splitting function of Classification and Regression Tree (CART). In particular, we propose two novel actuarial modified loss functions, namely, weighted sum of squared error and Canberra loss functions. These modified loss functions impose a significant penalty on grouping nonzero observations with zero ones at the splitting procedure, thus significantly improving separation of zeros and nonzeroes. Finally, we examine and compare, with a synthetic dataset and real dataset, the predictive performance of such actuarial modified tree-based models in relation to the traditional models. The results show that such modification leads to completely different tree structures and improved prediction performance.

6:25 PM – 6:50 PM
Title: Markov Chain Monte Carlo for Inference on Phase-Type Aging Model
Presenter: Cong Nie, Western University
Abstract: Bayesian inference for the Phase-Type Aging Model (PTAM) is considered when data consist of absorption times. Methodological extensions are developed based on an existing Markov chain Monte Carlo (MCMC) inference method. The extended methodology allows sampling from a multivariate posterior distribution, which complicates matters after data augmentation in the existing method. In this project, Gibbs sampling is developed to facilitate posterior sampling after data augmentation, which produces a two-level nested MCMC methodology. The outer level aims at data augmentation as usual, and the inner level samples from the posterior distribution via Gibbs sampling. The proposed approach is validated via a simulation study, wherein the parameters are estimated on the basis of simulated data from an underlying PTAM. Although our methodology is currently applied to PTAM, it is generally applicable to situations when the posterior distribution is not simplified enough after data augmentation.

The current bottleneck is the inefficient time of implementation, caused by the computation of large matrix exponentials. However, this may not be the case in future as the computational power is increasing at a steady pace. On the other hand, since matrix exponential is a unique characteristic of PTAM, we are considering avoiding this impediment by investigating the performance of our methodology in connection with other processes. We are also seeking other methodologies other than Gibbs sampling to sample from a multivariate posterior distribution.

Certain results on Bayesian inference in connection with a Discrete Multivariate Phase-Type Model will possibly be presented as well.

6:50 PM – 7:15 PM
Title: Structural varying-coefficient regression to differentiate individual policyholder's risk
Presenter: Rakheon Kim, Texas A&M University
Abstract: As new types of data such as personal driving records become available in the insurance industry, the needs for understanding policyholders' risk at individual level is growing. For example, the use of genetic testing in health insurance and the development of usage-based motor insurance where the premium depends on the driving records of policyholders require insurers to assess each policyholder's risk differently. However, classical actuarial models are based on the pooling of policyholders and do not account for the traits of individual policyholders sufficiently. In this study, our interest is to identify variables which determine individual policyholder's risk. This can be cast as a model selection problem for a varying-coefficient regression. However, this is challenging when there is a pre-specified group structure among variables. We propose a novel variable selection method for a varying-coefficient regression with such structured variables and provide a publicly available R package "svreg" for implementation of our method. Our method is empirically shown to select relevant variables consistently. Also, our method screens irrelevant variables better than existing methods. Hence, our method leads to a model with higher sensitivity, lower false discovery rate and higher prediction accuracy than the existing methods.

7:15 PM - 7:40 PM
Title: Natural Language Processing Application in Actuarial Science: Interpretable Language Models
Presenter: Zhiyu Quan, University of Illinois at Urbana-Champaign
Abstract: With the rapid emergence of InsurTech, a combination of cutting-edge technology applications and insurance operations, many traditional insurance companies are looking to further investigate alternative data, especially from social media and other online content, to stay competitive. In this paper, we apply interpretable natural language processing (NLP) techniques to turn raw unstructured text information into useful structured data for actuarial

analysis. Based on reviews and website contents provided by the InsurTech industrial partner, we performed sentiment analysis and topic modeling to create unbiased risk factors and refine business segmentation. This alternative actuarial-related information empowered by Insurtech innovations enhances the current rating factors for business insurance and further provides new angles to assess the underlying risk.

August 21, 2021

10:05 AM – 11:20 AM

Parallel Session I – Insurance Technology

10:05 AM – 10: 30 AM
Title: Machine learning (ML) applications in insurance
Presenter: Arnold Shapiro, Penn State University
Abstract: The focus of this presentation is on machine learning (ML) applications in insurance. To this end, the categories of ML algorithms that will be addressed are supervised learning based on regression and classification, unsupervised learning based on clustering and dimensionality reduction, and reinforcement learning. In each instance, an overview of insurance articles that apply the ML algorithms will be discussed. The presentation will end with a prognosis with respect to ML applications in insurance.

10:30 AM – 10:55 AM
Title: Leveraging High-Resolution Weather Information to Predict Hail Damage Claims: A Spatial Point Process for Replicated Point Patterns
Presenter: Lisa Gao, University of Wisconsin-Madison
Abstract: Technological advances in weather data collection allow insurers to incorporate high-resolution data to manage hail risk more effectively, but challenges arise when the response variable and predictors are collected from different locations. We propose a spatial mixed-effects framework for replicated point patterns to model the frequency and geographical distribution of hail damage claims following a hailstorm. Our model simultaneously incorporates traditional property rating characteristics collected from policyholders, as well as densely collected weather features, even when observed at different sets of locations across a region. We demonstrate inference and prediction using policyholder-level claims data from a U.S. insurer, supplemented with hail radar data and other spatially varying weather features. Incorporating rich, high-dimensional data to model the development of claim reporting patterns helps insurers anticipate and manage claims more efficiently.

10:55 AM – 11:20 AM

Title: Peer-to-peer multi-risk insurance and mutualaid plan
Presenter: Samal Abdikerimova, University of Illinois at Urbana-Champaign
Abstract: Peer-to-peer insurance is a decentralized network in which participants pool their resources together to compensate those who suffer losses. The rise of P2P insurance in Western countries, such as Friendsurance and Lemonade, has been viewed as disruptor to the traditional insurance industry in the same way Uber is to the taxi industry. A similar business model of mutual aid, such as the model developed by Xiang Hu Bao, has become popular in the East. It is a model which is designed to provide financial support to those in need and spreads the cost among all participants.

Despite the fast-changing landscape in this field, there has been scarce literature on the theoretical underpinning of P2P insurance and mutual aid. This paper presents the first effort to build a unified framework to quantify and assess the exchange of risks in a network of participants from different risk classes. Under this framework, the paper aims to address several essential research questions, including the fair exchange of heterogeneous risks and the optimality of algorithmic designs under various criteria. The modeling of multi-risk exchange is done with a P2P network structures. We show that these network structures not only can be used to explain the current practice in the industry, but also provides new tools to develop better designs. The paper concludes with a comparison of traditional insurance and P2P risk sharing from the standpoint of stability and cost reduction.

Parallel Session II – Data Science and Applications

10:05 AM – 10:30 AM
Title: Two-Stage Clustering to Assess the Impact of Low Dental Expenditures on Outcomes for Children
Presenter: Margie Rosenberg, University of Wisconsin-Madison
Abstract: Healthy People 2030 children's oral health goals recognize the importance on overall health and self-esteem in eliminating health disparities, and achieving health equity for children in their adult years. In this work, we focus on children (and their families) with low dental expenditures, and the potential impact on their overall and mental health, and behavioral issues at school or at home.

Using data from the calendar year 2016 Medical Expenditure Panel Study (MEPS), we form clusters using the Partitioning Around Medoids (PAM) algorithm to produce the first stage of the clustering process. PAM produces cluster centers representing actual individuals. The clusters are formed based on social determinants of the child and their dental services received (not expenditures or insurance), characteristics of the respondent to the survey, and household-level characteristics. The second stage clustering finds the pairwise correlation of the characteristic cluster percentages as input to a hierarchical clustering method to produce a

dendogram to pictorially reflect the similarities of the cluster profiles. We validate the work on an independent panel from the 2017 MEPS.

10:30 AM – 10:55 AM
Title: Patterns and Anomalies of Loss Development in P&C Insurance Market
Presenter: Steve Guo, Ball State University
Abstract: We analyze loss development in NAIC Schedule P loss triangles using functional data analysis methods. Relying on robust principal component analysis (RPCA), we study the incremental loss ratio curves of workers' compensation lines across hundreds of companies and 24 years. RPCA helps us to find out patterns of loss development, including (i) identifying outlier loss triangles; (ii) providing a dimension reduction tool to interpret the functional loss development data via a few factors. As one example of a relevant insight, we document distinctive loss development patterns between the late 1980s, 1990s and late 2000s periods. Moreover, our approach provides novel visualization tools. In the latter part of the article, we propose a functional model for generating probabilistic forecasts of incomplete cumulative loss ratio curves based on historical and similar development patterns.

10:55 AM – 11:20 AM
Title: Loss amount prediction using a double GLM with shrinkage and selection
Presenter: Gee Lee, Michigan State University
Abstract: The Gamma model has been widely utilized in a variety of fields, including actuarial science, where it has important applications in insurance loss predictions. Meanwhile, high dimensional models and their applications have become more common in the statistics literature in recent years. The availability of such high dimensional models have allowed the analysis of non-traditional data, including those containing textual descriptions of the response. In the models used in such applications, the dispersion may be designed to be related to a set of covariates, as opposed to being a single fixed value for the entire population. Following this approach, we incorporate a group Lasso type penalty in both the dispersion and the mean parameterization for a Gamma model, and illustrate its use in a predictive analytics application in actuarial science. In particular, we apply the method to an insurance claim prediction problem involving textual data analysis methods. Simulations are conducted to illustrate the variable selection and model fitting performance of our method.

Parallel Session III – Finance

10:05 AM-10:30 AM
Title: Model-Free Implied Dependence and the Cross-Section of Expected Stock Returns
Presenter: Yong Xie, University of Illinois Urbana Champaign

Abstract: Dependence in the market plays an important role in risk analytics, portfolio management, diversification, etc. In this paper, we provide new evidence of the importance of nonlinear characteristics in stock returns in the context of asset pricing.

Using comonotonicity theory, we develop a new, model-free measure of implied dependence (MFID). We first create a virtual comonotonic index as the counterpart to observed index in the comonotonic situation. We show that virtual European option with the virtual comonotonic index as underlying can be replicated by a portfolio of index component options with special strikes. Our MFID is defined as the ratio of model-free implied volatilities of index option and virtual comonotonic index option. Analogous to Pearson correlation, it is designed to gauge the distance between observed situation to the extreme comonotonic situation. Contrary to existing dependence measures, MFID captures both linear and nonlinear aspects in the return dependence structure, and thus provides a parsimonious quantification for overall market comovement.

We explore the implication of nonlinear dependence on asset pricing. Using Dow Jones Industrial Average index as a market index, the sort-portfolio trading strategy exploiting priced implied dependence risk generates significant abnormal returns. More specifically, we find stocks with high exposure to innovation in MFID deliver low returns cross-sectionally. The abnormal returns cannot be explained by linear correlation, implying existence of nonlinear dependence premium on the top of correlation premium. The dependence premium is robust to several empirical setups, and statistically significant after controlling standard risk factors such as Fama-French 5 factors, investment, liquidity and profitability.

10:30 AM – 10:55 AM
Title: The Hurst roughness exponent and its model-free estimation
Presenter: Xiyue Han, University of Waterloo
Abstract: We introduce the concept of the Hurst roughness exponent H as a generalization of the Hurst parameter to arbitrary continuous functions. The Hurst roughness exponent characterizes the regularity of a function in the following sense; a function admitting the Hurst roughness exponent H has vanishing p^{th} variation for all $p > 1/H$ and infinite p^{th} variation for $p < 1/H$. Moreover, we provide a class of model-free consistent estimators for the Hurst roughness exponent under simple assumptions. These estimators are model-free in the sense that they do not require any probabilistic assumptions and work with single time series. As an application, we apply our estimators to measure the roughness of the stock price and the realized volatility of various financial indices.

10:55 AM – 11:20 AM
Title: Valuing Lifetime Withdrawal Guarantees in RILAs
Presenter: Chenxin Xu, Temple University

Abstract: Registered index-linked annuities (RILAs) are equity-linked annuities offered in the United States that have seen a rapid rise in popularity in recent years, with sales increasing steadily to $24 billion in 2020. RILAs combine features from traditional variable annuities (TVAs) and fixed-index annuities (FIAs). They credit the investor periodic returns based on a popular stock index, subject to downside protection and upside limits (e.g. a 10% floor with a 15% cap over a 1-year crediting term).

Recently, some insurers have started to also offer Guaranteed Lifetime Withdrawal Benefits (GLWBs) as optional riders. And while the guarantees embedded in the RILA crediting mechanism are short-term and can be easily and near-perfectly hedged directly through the financial markets, the GLWB riders cannot. There is a large academic literature on pricing and hedging GLWBs embedded in TVA policies. However, RILA account values evolve very differently from TVA accounts, which will likely impact the GLWB value and risk exposure.

Therefore, in this study we explore the valuation of GLWB riders embedded in RILA products. We develop a model to determine the fair fee rate of such a rider and conduct a numerical analysis where we compare results across different crediting mechanisms as well as to TVAs and FIAs. Our preliminary results indicate that GLWB riders are much less costly when embedded in RILAs, compared to their TVA counterparts. We also find that they are less sensitive to the performance of the underlying index (i.e., they have a lower Delta), as a result of the short-term guarantees that impact the crediting of the periodic returns in a way that makes the RILA account value less volatile in the long run.

11:25 AM – 12:40 PM

Parallel Session I – Insurance Economics

11:25 AM – 11:50 AM
Title: Bowley vs. Pareto Optima in Reinsurance Contracting
Presenter: Mario Ghossoub, University of Waterloo
Abstract: The notion of a Bowley optimum has gained recent popularity as an equilibrium concept in problems of risk sharing and optimal reinsurance. In this paper, we examine the relationship between Bowley optimality and Pareto efficiency in a problem of optimal reinsurance, under fairly general preferences. Specifically, while we show that Bowley optimal contracts are indeed Pareto efficient (hence providing a first welfare theorem), we also show that only those Pareto efficient contracts that make the insurer indifferent between suffering the loss and entering into the reinsurance contract are Bowley optimal (hence providing only a partial second welfare theorem). We interpret the latter result as indicative of the limitations of Bowley optimality as an equilibrium concept in this literature. We also provide several illustrative examples.

11: 50 AM – 12:15 PM
Title: S-shaped narrow framing, skewness and the demand for insurance
Presenter: Shengchao Zhuang, University of Nebraska Lincoln
Abstract: The existing literature in insurance economics has shown that narrow framing can explain why people buy too little insurance compared to what standard theory predicts. However, there is also ample evidence suggesting people sometimes buy too much insurance. In this paper, we assume S-shaped narrow framing, i.e., the local utility function for evaluating the net insurance payoff is convex in the loss domain but concave in the gain domain, and show that it can reconcile with both insurance puzzles simultaneously. Especially, we show the policyholder under S-shaped narrow framing is more likely to underinsure more negatively skewed risks of loss but to overinsure less negatively skewed risks of loss. We further characterize the optimal insurance scheme under S-shaped narrow framing while incentive compatibility is satisfied. It contains a straight deductible when the net insurance payoff is negative but partial insurance when the net insurance payoff is positive.

12:15 PM – 12:40 PM
Title: Pseudo-Model-Free Hedging for Variable Annuities via Deep Reinforcement Learning
Presenter: Yuxuan Li, University of Illinois at Urbana-Champaign
Abstract: This talk applies a deep reinforcement learning approach to revisit the hedging problem of variable annuities. Instead of assuming actuarial and financial dual-market model a priori, the reinforcement learning agent learns how to hedge by collecting anchor-hedging reward signals through interactions with the market. By the recently advanced proximal policy optimization, the pseudo-model-free reinforcement learning agent performs equally well as the correct Delta, while outperforms the misspecified Deltas. The reinforcement learning agent is also integrated with online learning to demonstrate its full adaptive capability to the market.

Parallel Session II – Risk Measures

11:25 AM – 11:50 AM
Title: Reducing the runtime of Least Squares Monte Carlo in Risk Management
Presenter: Lu Xiong, Middle Tennessee State University
Abstract: The Least Squares Monte Carlo (LSMC) method was first proposed by Longstaff and Schwartz to price the American option, since then it has been applied in different industries from banking to the energy sector. In the last decade, there is an increasing demand for sophisticated risk modeling. To over the computational complexity of multi-risk factor LSMC, we propose distributed regression for LSMC. The idea of the distributed regression is fairly simple: instead of running the regression on one computer, we distribute the regression task to multiple computers (usually using cloud computers), then average the regressed coefficients to get the final regression equation. In this way, the computing time can be significantly reduced.

We can mathematically prove this simple idea can obtain the optimal regression results. GPU computing will also be explored on reducing its runtime.

11:50 AM – 12: 15PM
Title: Precise Large Deviations of Aggregate Claims with Arbitrary Dependence between Claim Sizes and Waiting Times
Presenter: Yiqing Chen, Drake University
Abstract: Consider a renewal risk model in which claim sizes and interarrival times correspondingly form a sequence of independent, identically distributed, and nonnegative random pairs with a generic pair (X, ?). Chen and Yuen (2012) studied precise large deviations of aggregate claims in this model under the assumption that (X, ?) obeys a dependence structure described via a stochastic boundedness condition on the waiting time ? for a large claim X. That assumption unfortunately leads to asymptotic independence between X and ? and hence considerably limits the usefulness of the result obtained there. In this short paper, we make an effort to avoid that assumption by allowing X and ? to be arbitrarily dependent. As by-products, we propose two novel applications of the main result, one to pricing insurance futures and the other to approximating both the value at risk and expected shortfall of aggregate claims.

12:15 PM – 12:40 PM
Title: Robust Estimation of Parameters for Log-Normal Insurance Payments with Data Truncation and Censoring
Presenter: Qian Zhao, Robert Morris University
Abstract: Log-normal distributions are useful tools for pricing insurance risks, modeling virus spread in public health, and analyzing the security of engineering system and in many other areas of application. In this paper, we develop a new estimation procedure - method of Winsorized moments (MWM) - to derive the parameters of log-normal insurance payments for which the loss data are affected by various loss control mechanism, e.g., truncation (due to deductibles), censoring (due to policy limits), and scaling (due to coinsurance factors). The asymptotic properties of the new estimators are provided and corroborated through simulations, and their performance (both robustness and computational efficiency) is compared to that of the method of trimmed moments (MTM) and the maximum likelihood estimators (MLE). Besides, we study the sensitivity of all estimators to an outlier that is placed at various locations ranging from the point of left- truncation to the point of right-censoring. The effect of model choice and parameter estimation method on risk pricing is also illustrated using actual data that represent 1500 Indemnity Losses in the US.

Parallel Session III – Data Science and Applications

11:25 AM – 11:50 AM

Title: A nonparametric sequential learning procedure for estimating the pure premium
Presenter: Liang Hong, The University of Texas at Dallas
Abstract: With the advent of the ``big'' data era, large-sample properties of a statistical learning method are becoming more and more important in an actuary's daily work. For a fixed sample size, regardless of how large it is, the variance of an estimator can be larger than a pre-assigned level to an arbitrary extent. In this paper, we propose a nonparametric sequential learning procedure for estimating the pure premium. Our method not only provides an accurate estimate of the pure premium but also guarantees that the mean of our random sample sizes is close to the unobservable optimal fixed sample size and the variance of our estimator is close to all small pre-determined levels. In addition, our method is nonparametric and applicable to any claims distribution; hence it avoids potential issues associated with a parametric model such as model misspecification risk and the effect of selection.

11:50 AM – 12:15 PM
Title: Fixed and Random Tail Probabilistic Robust Parametric Estimation Methodology
Presenter: Chudamani Poudyal, University of Central Florida
Abstract: Actuaries are required to predict the cost of uncertain future claims. Insurance claims are usually non-negative number, for example, automobile or liability insurance. The actuary's challenge is to determine an appropriate probability distribution that reflects the claims that will arrive in the next policy period. There are three challenges when constructing a probabilistic model. They are (1) past data are often affected by the presence of deductibles, policy limits and coinsurance, (2) past data may contain outliers (unusual observations that can distort the results), and (3) a need to evaluate the accuracy of the forecast. The contemporary literature in this area treats the three items as separate problems, leaving the practitioner with an incomplete framework for model building and assessment. This presentation will provide a comprehensive framework for addressing all three of these challenges simultaneously mainly via three estimation procedures: (a) maximum likelihood estimator (MLE), (b) dynamic method of trimmed or winsorized moments -- fixed tail probabilities, and (c) novel method of truncated moments with some variants - random tail probabilities. Asymptotic distributional properties are established (along with extensive simulation studies) which are essential parts for statistical inferences.

12:15 PM – 12:40 PM
Title: Adstools - An open source R library for Actuarial Science
Presenter: Trang Tran, Massachusetts Mutual Life Insurance Company (MassMutual)
Abstract: With the evolution of technology and advanced analytics, we are able to streamline most of our reporting processes across areas and domains for reproducibility and efficiency. Unfortunately, we have not seen much effort in the actuarial science domain in general, and in the life insurance field in particular, to create a modular and reusable collection of the common functionalities used by actuaries on a daily basis. The most recent effort, the "expstudies"

package built in R, was depreciated mid 2020. With such motivations, we introduce adstools, which is short for Actuarial Data Science Tools. Adtools provides a modular and reusable library with the primary functions needed to conduct a life experience study, with certain general utility functions applicable to other business needs. In this demo, we will introduce the primary functions that adstools is capable of delivering, as well as the specific use cases. The library is expected to be officially released on CRAN later this year.