The 21st M.Sc. Day

DEPARTMENT OF
STATISTICAL & ACTUARIAL SCIENCES

July 24, 2025

Table of Contents

Schedule of Events

Location: North Campus Building - 284

Time	Poster #	Last Name	First Name	Degree	Area	Supervisor(s)
9:30-9:50		**MORNING COFFEE**				
9:50-10:00		**OPENING REMARKS**				
Session 1 10:00-10:50	1	Elwaraki	Mahmoud	MSc	FM	Dr. Lars Stentoft
	2	Gonzalez Eslava	Rodrigo Ernesto	MSc	FM	Dr. Cristián Bravo
	3	Guo	Yiwei	MSc	FM	Dr. Ricardas Zitikis/ Dr. Cristián Bravo
	4	Hu	Minjie	MSc	FM	Dr. Marcos Escobar-Anel
	5	Braun	Phillip	MSc	ST	Dr. Hristo Sendov
	6	Groff	William	MSc	ST	Dr. Wenqing He & Dr. Pingzhao Hu
	7	Jiang	Hailong	MSc	ST	Dr. Ricardas Zitikis/ Dr. Hao Yu & Dr. Reg Kulperger
10:50-11:10		**COFFEE BREAK**				
Session 2 11:10-12:00	8	Li	Chenyao	MSc	AS	Dr. Shu Li
	9	Lam	Felix Ka Hay	MSc	FM	Dr. Cristián Bravo
	10	Palaisti	Marina	MSc	FM	Dr. Ricardas Zitikis/ Dr. Ankush Agarwal
	11	Pan	Yige	MSc	FM	Dr. Marcos Escobar-Anel
	12	Li	Peize	MSc	ST	Dr. Ricardas Zitikis/ Dr. Hao Yu & Dr. Reg Kulperger
	13	Li	Wenxin	MSc	ST	Dr. Grace Yi
12:00-1:00		**PHOTO TAKING + LUNCH IN WSC 263 & THE GRAD LOUNGE WSC 277**				
Session 3 1:00-1:50	14	Liu	Fenglin	MSc	AS	Dr. Shu Li
	15	Phillips	Alexander	MSc	FM	Dr. Marcos Escobar-Anel
	16	Varela Lopez	Ana Karen	MSc	FM	Dr. Cristián Bravo
	17	Zeng	Lina	MSc	FM	Dr. Hao Yu & Dr. Reg Kulperger
	18	Zeng	Weihang	MSc	FM	Dr. Serge Provost
	19	Li	William	MSc	ST	Dr. Dave Stanford, Dr. Hao Yu & Dr. Reg Kulperger
	20	Li	Yiming	MSc	ST	Dr. Ricardas Zitikis/ Dr. Hristo Sendov
1:50-2:10		**COFFEE BREAK**				
Session 4 2:10-3:00	21	Zhang	Mocun	MSc	FM	Dr. Lars Stentoft
	22	Zheng	Haoyuan	MSc	FM	Dr. Marcos Escobar-Anel
	23	Zheng	Jiayi	MSc	FM	Dr. Matt Davison & Dr. Cristián Bravo
	24	Liu	Mingqiu	MSc	ST	Dr. Grace Yi & Dr. Camila de Souza
	25	Ma	Sicong	MSc	ST	Dr. Katsu Goda & Dr. Jiandong Ren
	26	Wang	Yihan	MSc	ST	Dr. Hristo Sendov
3:00-3:20		**CLOSING REMARKS**				

Titles and Abstracts

Mahmoud Elwaraki, FM, supervised by Dr. Lars Stentoft

Replicating Teaching Economics to the Machines: Option Pricing

The aim of this project is to replicate the paper "Teaching Economics to the Machines" Chen et al. (2023), which applies deep learning with transfer learning to option pricing. We generated a deep residual neural network to learn the financial theory behind the option prices using synthetic data. We have generated data following the Black Scholes option pricing model and applied a training scheme similar to the paper. The training consisted of pretraining on a large dataset and then fine tuning some parts of the model on a smaller one. The smaller dataset was made to be a more realistic dataset. The final model predicts the option price, delta and vega and gives a MAE of the price and delta of 0.8756 and 0.0757, respectively. Feature importance analysis shows volatility and strike price as the most influential predictors. The model was also compared to one without the training steps in the report.

Rodrigo Ernesto Gonzalez Eslava, FM, supervised by Dr. Cristián Bravo

A Statistical Framework for Forward-Looking Probability of Default Estimation under IFRS 9 Using Vasicek-Based Macroeconomic Conditioning

IFRS 9 sets the accounting standard for how financial institutions recognize credit losses on financial assets. The framework involves calculating the Expected Credit Loss (ECL) based on three components: Probability of Default (PD), Loss Given Default (LGD), and Exposure at Default (EAD). Unlike Basel capital models, IFRS 9 requires a point-in-time (PIT) approach that incorporates both current and forecasted macroeconomic conditions. This project is oriented specifically on modelling forward-looking PIT PDs in compliance with IFRS 9. The methodology focuses on the integration of macroeconomic expectations into credit risk modelling using a two-stage process. In the first stage, observed default rates are transformed into Z-scores and regressed on macroeconomic drivers such as real GDP growth and unemployment. The models are selected based on a mix of statistical performance measures (R^2, AIC) and validated using diagnostic tests to ensure the assumptions of linear regression are met. In the second stage, a Vasicek single-factor model is applied to decondition baseline scorecard PDs and recondition them using macroeconomic forecasts, generating a full-term structure of PDs. Asset correlation (ρ) is empirically estimated using variance-based calibration. The framework is applied to three retail portfolios—mortgages, unsecured loans, and secured loans—producing responsive PD estimates that support IFRS 9-compliant ECL calculations.

Yiwei Guo, FM, supervised by Dr. Ricardas Zitikis/ Dr. Cristián Bravo

An Asymmetric Profit-Based Evaluation Method for Loss Given Default Models

Traditional accuracy metrics such as MSE or MAE treat over- and under-estimation of Loss-Given-Default (LGD) symmetrically, ignoring the fact that banks face markedly different economic consequences: over-provision ties up costly capital, whereas under-provision triggers fines and

retrospective capital charges. This project develops a fully monetary, asymmetric evaluation method for LGD models. We first derive a piece-wise linear loss function that multiplies each prediction error by exposure-at-default (EAD) and two cost coefficients—α for excess provisions and β (> α) for shortfalls—thereby "monetising" model errors during training. The same cost framework feeds into a per-loan profit formula and a threshold-based decision rule that yields an Expected-Maximum-Profit (EMP) cut-off for new-loan approval. Implemented as a custom objective in LightGBM and XGBoost, the method is back-tested on a retail-credit dataset and benchmarked against MSE-optimised models. Sensitivity analysis on α, β, and portfolio ROI quantifies robustness. The outcome is a deployable tool that aligns model evaluation with real bank profitability and regulatory capital management.

Minjie Hu, FM, supervised by Dr. Marcos Escobar-Anel

Comparative Analysis of Traditional and Dynamic Mean-Variance Optimization under Real and Simulated Market Conditions

This study conducts a comparative analysis of traditional and dynamic Mean-Variance Optimization (MVO) strategies for portfolio construction, evaluated under both historical and simulated asset return scenarios. Using six assets spanning a ten-year period from 2014 to 2024, we assess the performance of fixed-weight portfolios trained on early-period data and dynamic strategies which continuously updates mean return and covariance estimates using all available return data up to each rebalancing point. Four approaches are considered: traditional and dynamic MVO with daily and quarterly rebalancing. Performance is evaluated using annualized return, volatility, Sharpe ratio, and other performance measures. Results seem to indicate that quarterly outperforms daily rebalancing, with traditional quarterly strategies yielding higher Sharpe ratios. While dynamic MVO exhibits more stability in volatile or shifting environments, highlighting their adaptive advantage, it suffers from substantial estimation noise. Simulated analyses reveal both robust performance and potential extreme failures, underscoring the sensitivity of MVO strategies to input estimation and market regime changes. Sample size for estimations and other subsets of assets will be explored for robustness.

Phillip Braun, ST, supervised by Dr. Hristo Sendov

The Asymptotic Behaviour of Reals Roots in Random Polynomials

We perform a comprehensive analysis of two main theorems from the work "On Distribution of Zeros of Random Polynomials in Complex Plane" by Ibragimov and Zaporozhets. These theorems state that the proportion of real zeros of a wide class of random polynomials vanishes almost surely as the degree of the polynomial approaches infinity. The theorems also give a description of the behaviour for the expected number of real roots as the degree of the polynomial increases. Using probabilistic tools such as the Budan-Fourier theorem, the Kolmogorov-Rogozin inequality, and concentration function estimates, we derive the asymptotic sparsity of real zeros of polynomials with independent real-valued coefficients drawn from a finite set of non-degenerate distributions. The analysis includes derivations of supporting lemmas and bounds, large deviation estimates via Bernstein-type inequalities, and simulations of the theoretical results. This exposition aims to make these theorems more accessible by

expanding on the original proofs, and situating it within the broader landscape of universality in random polynomials.

William Robert Mcgaw Groff, ST, supervised by Dr. Wenqing He & Dr. Pingzhao Hu

Improving Microbiome-Based Disease Outcome Predictions With Conformal Inference

Due to the highly complex nature of microbiome data, cutting-edge models are typically complicated black box neural networks for which traditional inference techniques are not available. In high-stakes classification settings where we are aiming to diagnose patients based in part on their gut microbiome, being unable to quantify the uncertainty of our predictions can be problematic. In this project, we make use of conformal prediction to generate prediction sets with theoretical coverage guarantees without needing to make restrictive modelling assumptions. We aim to adapt the microbiome modelling architecture proposed by Fung et al. (2023), which uses a hybrid Convolution Neural Network Long Short-Term Memory model approach, to include several additional conformation prediction techniques. The resulting conformal predictors were evaluated on the DIABIMMUNE and PROTECT datasets. We found that cross-conformal predictors tended to produce more efficient prediction sets, whereas split-conformal predictors tended to more closely follow the desired confidence level.

Hailong Jiang, ST, supervised by Dr. Ricardas Zitikis/ Dr. Hao Yu & Dr. Reg Kulperger

Geometric Boundary Detection Approach for Cold-Start Active Learning: A Theoretical Framework and Experimental Validation

Addressing the sample selection problem in cold-start active learning scenarios, this study presents a novel method termed B-gradient. The proposed approach identifies cluster boundaries by exploiting the geometric asymmetry of data point neighborhoods and employs gradient computation as a measure of information gain. Specifically, the method maps each data point's k-nearest neighbors onto a unit sphere, computes the relative magnitude of their sum vectors, and calculates gradients with respect to the farthest neighbor. This design eliminates the requirement for prior knowledge while demonstrating robustness against outliers and class imbalance. Experimental validation on benchmark datasets, including ECG5000, reveals that B-gradient achieves competitive overall accuracy while delivering superior class-balanced performance compared to baseline approaches.

Chenyao Li, AS, supervised by Dr. Shu Li

The Choice of Rehabilitation Barrier under Chapter 11 Bankruptcy Model

This project investigates the optimal selection of the rehabilitation barrier within a contemporary regulatory framework. In traditional ruin theory, bankruptcy is considered immediately when the firm's surplus process drops below a pre-determined barrier, say level 0. However, another option is to allow the firm to reorganize and financially recover during a grace period. In practice, these two bankruptcy rules align with the Chapter 7 liquidation and Chapter 11 reorganization of the U.S. Bankruptcy Code.

We focus on the three-barrier model for Chapter 11 bankruptcy framework, which consists of the (lower) liquidation barrier, the (upper) recovery barrier, and the rehabilitation barrier. Our study specifically examines how a firm may strategically choose the rehabilitation barrier to minimize the probability of liquidation. We derive a closed-form expression for the total liquidation probability. Numerical experiments with both synthetic parameters and real-world surplus data demonstrate the existence of an interior optimal rehabilitation barrier, which balances early intervention benefits against the cost of premature filing. The findings offer practical implications for risk management in the insurance sector.

Felix Ka Hay Lam, FM, supervised by Dr. Cristián Bravo

Transformer-Based Deep Learning Models for Mortgage Credit Scoring

This study investigates the application of transformer-based deep learning models specialized in tabular data for credit risk assessment at the mortgage application stage. Utilizing a structured dataset derived from Fannie Mae Single-Family Loan Performance data, the objective is to predict binary default outcomes and develop an effective application scorecard. After thorough data selection and preprocessing, we implement and evaluate conventional methods such as Logistic Regression and XGBoost, and modern deep learning model architectures designed for tabular data, including SAINT and the FT-Transformer. Model performance is evaluated using the area under the receiver operating characteristic curve (AUC), while the interpretability of each modeling approach is assessed through SHAP value analysis. The results demonstrate that tabular transformer-based deep learning models can offer competitive, and in some cases superior, performance over traditional methods, highlighting their potential for enhancing credit scoring practices in the mortgage industry.

Marina Palaisti, FM, supervised by Dr. Ricardas Zitikis/ Dr. Ankush Agarwal

Climate-Aware Capital Allocation in Loan Portfolios under NGFS scenarios

This paper develops a forward-looking methodology for capital allocation in large loan portfolios exposed to climate risk over a 3–5 year horizon. Firm-level production is modeled as a stochastic process with scenario-adjusted drift, reflecting economic impacts from NGFS 2024 short-term climate scenarios across five distinct transition and physical risk pathways. These dynamics influence a nonlinear value process, which determines obligor default. Joint default simulations are performed via Monte Carlo, with dimensionality reduction achieved through Principal Component Analysis (PCA). The resulting loss distributions support a capital allocation strategy optimized via linear programming, subject to portfolio-wide constraints. Our findings illustrate how capital needs vary significantly across climate scenarios, emphasizing the importance of incorporating scenario-based stress testing into Internal Capital Adequacy Assessment Process (ICAAP) frameworks. The proposed approach is tractable, interpretable, and adaptable to regulatory capital planning.

Yige Pan, FM, supervised by Dr. Marcos Escobar-Anel

Parameter Recovery of a Gaussian Affine Discrete-Time Interest-Rate Model

This report studies the recovery of parameters for a newly developed Gaussian affine discrete-time interest-rate model, assuming the interest rates path is known. The model has five parameters, but our focus is on the main three parameters: the long-run driver α, the autocorrelation β, and the volatility γ. Synthetic bond prices with maturities of 1, 5, 10, 30 years are generated using the model's closed-form pricing formulas. Our objective function measures the error in pricing. We use a two-stage procedure combining grid search and local optimization. Results show that the two-parameter model (α, β) and the three-parameter model (including γ) can be robustly recovered, with up to 0.1% error across a wide range of initial values. Finding the true parameters seems elusive, particularly for three and the upcoming four parameters case due to the roughness of the objective function.

Peize Li, ST, supervised by Dr. Ricardas Zitikis/ Dr. Hao Yu & Dr. Reg Kulperger

Forecasting Ontario's Daily Maximum Electricity Supply with Weekly Memory and Maximum Temperature

We analyze Ontario's daily maximum electricity supply from 2015 to 2024 and its link to maximum daily temperature. A two-segment temperature term separates cold- and warm-weather demand; weekly memory is captured with one-, seven- and fourteen-day lags of past peaks. Trained on the full ten-year data set, the model explains 79 % of historical variation. Applied to an out-of-sample period (1 Jan – 31 May 2025) it achieves MAE = 820 MW, R^2 = 0.67 and 90 % interval coverage = 0.85, outperforming polynomial, LOESS and tree-based benchmarks while using only five coefficients. Weak residual autocorrelation and heavy tails remain, but heteroskedasticity is negligible, so Newey-West errors provide reliable inference. The compact structure makes the model suitable for day-ahead operations and long-range planning.

Wenxin Li, ST, supervised by Dr. Grace Yi

An R Package for Feature Screening with Large-Scale and High-Dimensional Survival Data

Survival analysis has attracted extensive research attention over the past five decades, following the seminal papers by Cox (1972; 1975) on the proportional hazards model. While numerous methods have been proposed to address various features of survival data, research gaps remain, particularly in settings involving high-dimensional data. In contrast to many studies that consider the high-dimensional scenarios with a large covariate dimension *p* and small sample size *n*, Yi, He, and Carroll (2022) developed a model-based feature screening method tailored for large-scale survival data where *n* is large and *p* grows faster than a polynomial in *n* (i.e., the so-called NP-dimension). Their method is developed by leveraging the unique connection between the Cox proportional hazards model and logistic regression to achieve computationally efficient exclusion of unimportant covariates. Although Yi, He, and Carroll (2022) provided theoretical justification for their method, its practical use has been limited by the lack of a publicly available software package. In this study, we implement their method and introduce an R package designed to make it accessible to applied analysts.

Fenglin Liu, AS, supervised by Dr. Shu Li

Phased Retirement Decision Making in a Deterministic Economy

Phased retirement is getting more popular for retirement planning, thanks to its flexibility in offering a smoother transition, continued engagement, and greater control over how and when to retire. This project investigates the optimal phased retirement decisions within a deterministic economy. Building on the model of Chen, Hentschel, and Steffensen (2021), we formulate the problem as the optimization over the individual's expected utility of lifetime consumption, where a constant relative risk aversion utility function is assumed. In the presence of phased retirement, we aim to address two key decisions on the timing and duration of phased retirement. Due to the complexity of the model, closed-form analytical solutions are generally not attainable, even in a deterministic economic setting. Consequently, we adopt numerical methods and simulation techniques to explore the optimal phased retirement strategies. Our results indicate the potential benefits of incorporating phased retirement into the retirement planning.

Alexander Phillips, FM, supervised by Dr. Marcos Escobar-Anel

A Theoretical and Empirical Analysis of Variational Autoencoders: An Application in Finance

A Variational Autoencoder (VAE) is a type of generative model that learns the joint distribution of data and its underlying latent space. It typically employs a standard normal distribution as the latent prior regularizer. By encoding data into a structured latent space and decoding samples drawn from it, the VAE enables the generation of new data points that resemble those from the original distribution. We prove that even for a shallow neural network with a one-dimensional encoder and latent space, the number of optimal solutions depends on the statistical properties of the input data, with a unique solution guaranteed only in specific cases. We also demonstrate the model's inability to capture fat-tailed distributions, as the standard normal prior decays too rapidly to adequately represent the low-density regions of such distributions.

Ana Karen Varela Lopez, FM, supervised by Dr. Cristián Bravo

Automated Risk-Based Decisioning Framework for Pre-Approved Credit Offers

This project delivers an automated risk decision-making framework aimed at evaluating existing customers for new product pre-approvals or credit line increases. The model streamlines the credit decision process by automatically assigning approved or denied outcomes. This automation enhances decision consistency, reduces underwriting intervention, and increases confidence across the sales and credit teams. The solution is built using a layered structure that combines six strategic risk overlays with two predictive score models. These components work together to assess both the customer's credit risk profile and financial capacity, enabling tailored credit offers that align with the institution's risk appetite and credit policies. As part of its deployment, the model includes a detailed evaluation of swap-in/swap-out populations relative to the current decision-making model. This ensures that any shifts in approval or rejection volumes are well understood and risk-aligned. The design is also

synchronized with recent updates across broader credit strategies, ensuring consistency and cohesion across the decision-making framework.

Lina Zeng, FM, supervised by Dr. Hao Yu & Dr. Reg Kulperger

Time-Series and Categorical Analysis of IT Risk Events in a Financial Services Company

This study presents a data-driven approach to IT risk monitoring within the context of a global financial services organization. Leveraging over 200 risk event records extracted from Jira, a time-series analysis was conducted to examine the frequency, categorization, and severity of IT risk incidents over a multi-year period. Risk events were aggregated and visualized using Python-based tools to identify monthly trends, distribution by category, and severity levels. A cross-sectional analysis of event status and severity revealed key vulnerabilities, including a subset of critical risks that remained unresolved, highlighting areas requiring immediate strategic attention. The findings supported ongoing risk governance efforts and contributed to high-level discussions regarding IT strategy alignment. This work demonstrates the practical application of quantitative methods in enterprise risk management, and underscores the value of integrating operational data into long-term decision-making frameworks.

Weihang Zeng, FM, supervised by Dr. Serge Provost

Portfolio Construction via Log-Return Copula Analysis

This study proposes a portfolio construction methodology grounded in the structural dependence exhibited by asset pairs in the copula domain. Log-return transformations are applied to asset price series to remove deterministic trends and stabilize variance, thereby rendering the data approximately stationary and amenable to copula-based modeling. Particular emphasis is placed on detecting joint tail behavior through localized peaks in the estimated copula density of the log-returns. These tail dependencies serve as criteria for asset selection, where custom-developed count and vertex indices quantify the significance of such extremal clustering. While the methodology is presented as a complement to conventional investment strategies, it serves as an independent screening mechanism grounded in dependence structures rather than marginal characteristics. The analysis centers on the bivariate setting, motivated by the flexibility of vine copula constructions, which decompose higher-dimensional dependencies into bivariate components. Several illustrative examples are provided to demonstrate the practical applicability of the approach.

William Li, ST, supervised by Dr. Dave Stanford, Dr. Hao Yu & Dr. Reg Kulperger

Evaluating the Adequacy of the Compound Poisson process for Ontario Forest fires 1976 - 2024

Podur et al. (2009) proposed a compound Poisson process to model the annual area burned by large forest fires in Ontario, assuming a constant-intensity Poisson process for fire counts and a Weibull distribution for fire sizes. In this study, we revisit and extend their framework using data from 2006 onward. First, we examine whether the key distributional moments for fire counts and fire sizes have shifted since 2005 by comparing the first, second, and third empirical moments to those reported in

Podur's work. Second, we evaluate the adequacy of the compound Poisson model through simulation and find that it underestimates the observed variability in annual area burned. To address this, we introduce a Cox–Poisson process in which fire count intensity varies stochastically across years according to a gamma distribution. This formulation induces a negative binomial distribution for fire counts and results in a more accurate reproduction of total area burned in simulation. We also confirm that the Weibull distribution remains a better fit for fire sizes compared to exponential and Pareto alternatives.

Yiming Li, ST, supervised by Dr. Ricardas Zitikis/ Dr. Hristo Sendov

A Survey of Some Classic Limit Theorems in Probability Theory

This project introduces some of the most important limit theorems in classical probability theory in an easy-to-understand way. These theorems explain the behavior of random events when they are repeated many times, revealing surprising laws behind randomness. The project focuses on four core results: the Strong Law of Large Numbers (SLLN), the Central Limit Theorem (CLT), the Law of the Iterated Logarithm (LIL), and the Berry-Essing Theorem. For each theorem, the project not only provides its standard form, but also discusses different versions. For example, the project explores the Strong Law of Large Numbers under pairwise independence (rather than perfect independence), and the Central Limit Theorem when the variables are not identically distributed. These differences show how these theorems still hold under more relaxed or more realistic assumptions. This project also discusses two fundamental 0-1 laws (Kolmogorov and Hewitt Savage), which show that certain events are either almost certain to occur or almost certain not to occur. At the same time, all theorems in the article provide detailed proofs and intuitive explanations.

Mocun Zhang, FM, supervised by Dr. Lars Stentoft

Multilevel Monte Carlo for Cost-Aware Dynamic Portfolio Optimization

This project discusses the application of Multi-level Monte Carlo (MLMC) to enhance dynamic portfolio optimization, which incorporates trading costs. Two strategies were tested; one in which costs are ignored, i.e. cost-blind trading, and one in which only one-step fees are considered in trading, i.e. one-step trading. MLMC improves estimations by performing simulations with varying levels of refinement, such as coarse crude approximations and fine-grained detailed approximations side by side and then bundling these together. The key finding is that MLMC is very efficient in achieving accuracy: it significantly reduces the variance of estimated returns while reducing computational costs compared to standard Monte Carlo techniques to achieve the same level of accuracy. In the case of cost-blind strategies, MLMC cut down redundant trading by 12 percent and enhanced returns. In case of one-step strategies, it was consistent with little fluctuations. Such gains add to the effectiveness and precision of empirical dynamic portfolio management.

Haoyuan Zheng, FM, supervised by Dr. Marcos Escobar-Anel

Analytical Pricing of Insurance Products - A sensitivity study on surrender and withdrawal rates

This project studies the sensitivity of variable annuities (VAs), including guaranteed minimum accumulation benefits (GMABs), guaranteed minimum death benefits (GMDBs), and more recent related insurance products, to important model parameters. We implement, via Monte Carlo, the analytical framework developed by Escobar et al. (2016), incorporating stochastic modeling of financial markets, mortality risk, and policyholder surrender behavior. The numerical analyses conducted use real-world motivated parameters and simulated data. The primary goal is to evaluate the sensitivity of the prices to surrender and withdrawal rates; we also study sensitivity to market parameters, and explore the potential for optimal rates under practical modeling assumptions.

Jiayi Zheng, FM, supervised by Dr. Matt Davison & Dr. Cristián Bravo

How to Hedge Options on a Cross-Border Dual Listed Stock

Many publicly traded Canadian companies, such as major banks and mining firms, are cross-listed on both the Toronto Stock Exchange (TSX) and the New York Stock Exchange (NYSE), with their shares quoted in Canadian dollars (CAD) on the TSX and in U.S. dollars (USD) on the NYSE. This dual listing presents opportunities for option hedgers to decide on which market to hedge their portfolio. In this project, we study options on the Canadian dollar-denominated stock price R (as traded on the TSX) and on the cross-currency product R = F S, where S is the USD price of the same stock on the NYSE, and F is the CAD in USD foreign exchange rate. We derive the partial differential equations (PDEs) that govern the prices of both types of options: the option on R follows the standard Black-Scholes model, while the option on R = F S leads to a two dimensional variant of the Black-Scholes PDE. We compute analytic pricing solutions in both cases and derive the corresponding hedging strategies. Our goal is to investigate how different hedging choices are affected by factors such as differential transaction costs or alternative hedging strategies, and how these differences impact replication error and risk exposure in a realistic market environment.

Mingqiu Liu, ST, supervised by Dr. Grace Yi & Dr. Camila de Souza

Survival Analysis of Hospital Admissions Data with Missing Observations

This study investigates a hospital admissions dataset that contains missing observations. Our objective is to identify clinical factors associated with time to death. We employ a Cox regression model to analyze the relationship between in-hospital survival time and predictors, such as age, gender, ejection fraction (EF), brain natriuretic peptide (BNP), diabetes, and hypertension. We examine the impact of either ignoring missing data or accounting for it in the analysis. Specifically, we compare three approaches: complete-case analysis, K-nearest neighbours (KNN) imputation, and multiple imputation. The study suggests that older age, lower EF, higher BNP, diabetes, and hypertension are significant predictors of shorter survival.

Sicong Ma, ST, supervised by Dr. Katsu Goda & Dr. Jiandong Ren

Probabilistic Modeling and Parametric Insurance Design for Earthquake Losses in British Columbia

British Columbia (BC) faces significant seismic risk due to its proximity to the Cascadia subduction zone and the high exposure of infrastructure and population. Traditional indemnity-based earthquake insurance in BC suffers from high premiums, delayed claims processing, and administrative inefficiencies, particularly in the context of spatially correlated and hard-to-verify losses. As an alternative, parametric insurance offers pre-defined payouts triggered by physical parameters—such as earthquake magnitude and ground shaking intensity—typically reported by public agencies, without requiring post-disaster loss verification. However, the challenge of basis risk, defined as the mismatch between payouts and actual losses, limits its effectiveness. This study leverages simulated earthquake data at the Forward Sortation Area (FSA) level in BC and develops a logistic regression model to estimate the probability of insured losses exceeding certain threshold. By optimizing trigger thresholds for both the insured and the insurer, the study proposes a compromised solution that balances the interests of both parties. Building on recent literature, this research introduces a probabilistic framework for designing index-based parametric earthquake insurance in high-risk regions.

Yihan Wang, ST, supervised by Dr. Hristo Sendov

Theory and Numerical Simulations of the Asymptotic Distribution of Zeros of Complex Random Polynomials

In this work, we consider two of the main results from the paper by Ibrahimovic concerning the asymptotic distribution of zeros of random polynomials. Let $\{\xi_k\}$ be in dependent identically distributed real- or complex-valued random variables such that $\boldsymbol{E}\log(1+|\xi_0|) < \infty$, and consider the random polynomial $G_n(z) = \sum_{k=0}^{n} \xi_k z^k$. We show that this moment condition is both necessary and sufficient for the empirical distribution of the zeros to converge almost surely to the uniform distribution on the unit circle. Furthermore, we prove that the arguments of the zeros become asymptotically equidistributed in $[0,2\pi)$. The goal of this exposition is to explain these results in a clear and accessible way, by providing all the necessary background and mathematical details. The results are further supported by numerical simulations.