ISFA

GRANDE ÉCOLE
D'ACTUARIAT
ET DE GESTION
DES RISQUES

INSTITUT DE SCIENCE FINANCIÈRE ET D'ASSURANCES

♛ COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Université Claude Bernard Lyon 1

Université Claude Bernard (UCB) Lyon 1

Dear participant,

We would like to take the opportunity to thank you for joining us at the first Lyon-Columbia research workshop on Actuarial Science.

This conference, jointly organized by SAF research laboratory (ISFA, Université Lyon 1) and Columbia University (Stats and IEOR departments in particular), is an opportunity for practitioners and academics to participate to research discussions and interactions on Quantitative Risk Management and Data Science for Insurance and Finance.

These topics are at the core of our research in the context of the DAMI Chair (Data Analytics and Models for Insurance) and the LoLitA ANR project.

Over the next two days, Columbia and SAF researchers, as well as special guests, will present their research works and interact in special sessions to foster future collaborations.

This conference is also the opportunity to honor the great mathematician and political writer, Emil Julius GUMBEL (1891-1966), on the occasion of the 125th anniversary of his birth. In addition to his outstanding work in mathematics, including the Gumbel distribution and the development of extreme value theory, as well as his political commitments against nazism, a particular detail about Emil GUMBEL is that he taught at ISFA in the 30's and held a chair at Columbia University after he moved to the US.

We would like to thank his daughter-in-law, Patricia GUMBEL, for joining us from California for the occasion, as well as all the speakers of the conference and ceremony.

We wish you all a thrilling event with inspiring new insight.

José Blanchet   *Stéphane Loisel*

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

ISFA

HOST ....

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Columbia University (officially Columbia University in the City of New York) is a private, Ivy League, research university in Upper Manhattan, New York City. It was established in 1754 as King's College by royal charter of George II of Great Britain. Columbia is the oldest college in New York State and the fifth chartered institution of higher learning in the country, making it one of nine colonial colleges founded before the Declaration of Independence. After the revolutionary war, King's College briefly became a state entity, and was renamed Columbia College in 1784. A 1787 charter placed the institution under a private board of trustees before it was renamed Columbia University in 1896.

Columbia is one of the fourteen founding members of the Association of American Universities, and was the first school in the United States to grant the M.D. degree. The university is organized into twenty schools, including Columbia College, the School of Engineering and Applied Science, and the School of General Studies. Columbia annually administers the Pulitzer Prize. Notable alumni and former students (including those from King's College) include five Founding Fathers of the United States ; nine Justices of the United States Supreme Court ; 20 living billionaires ; 29 Academy Award winners ; and 29 heads of state, including three United States Presidents. Additionally, some 100 Nobel laureates have been affiliated with Columbia as students, faculty, or staff, second in the world only to Harvard.

Source : Wikipedia

www.columbia.edu

HOST

GRANDE ÉCOLE
D'ACTUARIAT
ET DE GESTION
DES RISQUES

ISFA INSTITUT DE SCIENCE FINANCIÈRE ET D'ASSURANCES

LABORATOIRE
SAF
SCIENCES ACTUARIELLE
& FINANCIÈRE

ISFA Graduate School of Actuarial Studies is an interdisciplinary education and research institute in actuarial, financial sciences and risk management.

Created in 1930, it is the oldest accreditated French actuarial program.

ISFA has developed actuarial programs overseas (in Vietnam, Lebanon, Morocco, Senegal), as well as a wide student exchange program. ISFA is also the leader of an innovative capacity building program: training actuarial professors in Africa and Latin America...

Graduate School of University Claude Bernard Lyon 1, ISFA welcomes over 600 students a year and hosts the only French research Lab dedicated to actuarial science: SAF Laboratory.

www.isfa.fr

The SAF (Actuarial Science and Finance) research laboratory undertakes interdisciplinary research on risks in insurance and finance:
- risk models and measures (probability and statistics)
- risk management and Enterprise Risk Management
- economic and financial risk analysis
- accounting and regulation in insurance and finance.

The research topics of the laboratory are always evolving to include new risks (human longevity improvements, natural hazards,…), recent accounting standards (IFRS 4 Phase 2), new prudential regulation systems (Basel 3, Solvency 2) as well as new risk management practices (Enterprise Risk Management). Current team research projects focus in particular on the impact of modeling and analytics on the management of insurance and financial firms, environmental risks, as well as longevity risk management.

isfa.univ-lyon1.fr/recherche/

ISFA-COLUMBIA

WORKSHOP PROGRAM
MONDAY JUNE 27, 2016 ....

8:30	*Welcome coffee*
9:15	**WELCOME ADDRESS**
	José Blanchet (Columbia University) & Stéphane Loisel (ISFA, SAF Laboratory University Lyon 1)
9:20	**SHORT PRESENTATION OF LOLITA PROJECT & ASSOCIATED BIG DATA CHALLENGES**
	Nicole El Karoui (Paris 6 University) & Stéphane Loisel (ISFA, Lyon 1 University)
9:30	**BIRTH MONTH AFFECTS LIFETIME DISEASE RISK: A PHENOME-WIDE METHOD**
	Nicholas Tatonetti (Columbia University)
10:20	**CANCER BASELINE AND OTHER BIG DATA APPROACHES TO EXTEND LIFE**
	Edouard Debonneuil (SAF, Lyon 1 University)
11:10	*Coffee break*
11:30	**APPLICATIONS OF DEEP LEARNING IN BIOMEDICINE, PART 1**
	Alexander Zhavoronkov (In Silico)
13:00	*Lunch break & discussions*
14:30	**STRUCTURING TEMPORAL SPARSE DATA WITH APPLICATION TO OPINION MINING**
	Julien Velcin (ERIC, University Lyon 2)
15:20	**TREE-BASED CENSORED REGRESSION WITH APPLICATIONS TO INSURANCE**
	Xavier Milhaud (ISFA, Lyon 1 University)
16:10	*Refreshment break*
16:30	**RESEARCH DISCUSSIONS WITH DAMI CHAIR TEAM**
	Chairs Christian Robert (ISFA, SAF Lyon 1 University) & Frédéric Planchet (ISFA, SAF Lyon 1 University)
	- Presentation of the research chair Data Analytics and Models in Insurance
	- Presentation of a research project by Christian Robert
	- Feedback on Cardif Kaggle competition by Lam Dang (BNP Parisbas Cardif)
	- Discussions
17:45	**RESEARCH DISCUSSIONS WITH LOLITA TEAM & COLUMBIA TEAM**
	Chairs Nicole El Karoui (Paris 6 University), Stéphane Loisel (ISFA, Lyon 1 University)
	& Olivier Lopez (Paris 6 University)

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK   ISFA

ISFA-COLUMBIA

TUESDAY JUNE 28, 2016

8:30   *Welcome coffee*

9:15   **LATENT AND NETWORK MODELS WITH APPLICATIONS TO FINANCE**
JingChen Liu (Columbia University)

10:05   *Coffee break*

10:30   **APPLICATIONS OF DEEP LEARNING IN BIOMEDICINE, PART 2**
Alexander Zhavoronkov (In Silico)

12:00   *Lunch break & discussions*

13:30   **DISTRIBUTIONAL ROBUSTNESS AND REGULARISATION IN MACHINE LEARNING:
TWO SIDES OF THE SAME COIN**
Karthyek Murthy (Columbia University)

14:20   **SRI INVESTMENT: APPLYING A SUPERVISED LEARNING ALGORITHM TO ENVIRONMENTAL,
SOCIAL AND GOVERNANCE HISTORICAL SCORES**
Christophe Geissler (Advestis & DAMI chair)

15:10   *Refreshment break*

15:40   **ON ROBUST RISK ANALYSIS**
José Blanchet (Columbia University)

16:30   **OVERREACTING RANDOM WALKS TEND TO BECOME AMNESIC**
Nabil Kazi-Tani (ISFA, SAF Lyon 1 University)

17:20   *Break*

17:45   **CEREMONY IN HONOUR OF E. J. GUMBEL**
- Emil J. GUMBEL - An Extreme Statistician by Matthias Scherer (University of Munich)
- Additional remarks by Officials from Lyon
- Official renaming of Lecture Hall G3 after the name of GUMBEL

19:00   *Cocktail*

BIRTH MONTH AFFECTS LIFETIME DISEASE RISK: A PHENOME-WIDE METHOD

Nicholas TATONETTI
Columbia University

Dr. Nicholas Tatonetti is assistant professor of biomedical informatics in the Departments of Biomedical Informatics, Systems Biology, and Medicine and is Director of Clinical Informatics at the Herbert Irving Comprehensive Cancer Center at Columbia University.
He received his PhD from Stanford University where he focused on the development of novel statistical and computational methods for observational data mining. He applied these methods to drug safety surveillance where he discovered and validated new drug effects and interactions.

His lab at Columbia is focused on expanding upon his previous work in detecting, explaining, and validating drug effects and drug interactions from large-scale observational data. Widely published in both clinical and bioinformatics, Dr. Tatonetti is passionate about the integration of hospital data (stored in the electronic health records) and high-dimensional biological data (captured using next-generation sequencing, high-throughput screening, and other «omics» technologies). Dr. Tatonetti has been featured by the New York Times, Genome Web, and Science Careers. His work has been picked up by the mainstream and scientific media and generated thousands of news articles.

An individual's birth month has a significant impact on the diseases they develop during their lifetime. Previous studies reveal relationships between birth month and several diseases including atherothrombosis, asthma, attention deficit hyperactivity disorder, and myopia, leaving most diseases completely unexplored. This retrospective population study systematically explores the relationship between seasonal affects at birth and lifetime disease risk for 1688 conditions. We developed a hypothesis-free method that minimizes publication and disease selection biases by systematically investigating disease-birth month patterns across all conditions. Our dataset includes 1 749 400 individuals with records at New York-Presbyterian/Columbia University Medical Center born between 1900 and 2000 inclusive. We modeled associations between birth month and 1688 diseases using logistic regression. Significance was tested using a chi-squared test with multiplicity correction. We found 55 diseases that were significantly dependent on birth month. Of these 19 were previously reported in the literature (P<.001), 20 were for conditions with close relationships to those reported, and 16 were previously unreported. We found distinct incidence patterns across disease categories. Lifetime disease risk is affected by birth month. Seasonally dependent early developmental mechanisms may play a role in increasing lifetime risk of disease.

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

ISFA

CANCER BASELINE AND OTHER BIG DATA APPROACHES TO EXTEND LIFE

Edouard DEBONNEUIL
SAF, Lyon 1 University
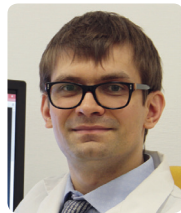
Edouard holds masters of engineering, biostatistics, bio-gerontology and actuarial risks from Centrale Paris, UCLA, Paris V, ISFA. He worked on web behavioral models (IBM Research Center in Germany, IPO.com in the USA), biomedical and bioinformatics research (cardiopulmonary remodelling at Marie-Lannelongue and Inserm, genetics at the Pasteur Institute), actuarial derivatives and statistics from large databases regarding financial risks, longevity risks and health risks (Derivatives at HSBC, AXA Group and AXA France, Celtipharm).

Passionnate in transforming societies towards longer and healthier lives, he tries to have an understanding both from biomedical and actuarial sciences and he is co-founder and member of the board of the International Longevity Alliance, an association present in more than 50 countries.

The era of big data brings an unprecedented promise for human health and longevity: we will be able to learn from the individual lives of billions of people, to better know what is good and bad and to pilot our health much more effectively. Here, such a feedback loop is already presented with agregated data about 4 billion persons -- a project called «Baseline», on which more than 200 persons have joined forces in the last 6 months -- and approaches with large-scale i ndividual data are introduced.

10:20

APPLICATIONS OF DEEP LEARNING IN BIOMEDICINE

Alexander ZHAVORONKOV
Insilico Medicine

11:30

Alex Zhavoronkov, PhD is the CEO of Insilico Medicine, Inc a Baltimore-based company utilizing big data analysis and deep learning for aging research and drug discovery. He also heads the International Aging Research Portfolio (IARP) knowledge management system for aging research and serves as the chief science officer of the Bio-gerontology Research Foundation in the UK. Prior to Insilico Medicine, he co-founded the First Oncology Research and Advisory Center (FO-RAC), served as the director of ATI Technologies (Nasdaq: AMD) and as the director of GTCBio.

Dr. Zhavoronkov is the author of over forty peer-reviewed scientific as well as popular papers and books including "The Ageless Generation: how biomedical advances will transform the global economy" published by Palgrave Macmillan.
He holds two bachelor degrees from Queen's University, a masters in biotechnology from Johns Hopkins University and a PhD in physics and mathematics from the Moscow State University and is the international adjunct professor at the Moscow Institute of Physics and Technology.

Evgeny PUTIN
Insilico Medicine

Evgeny Putin is a Deep Learning Lead at Insilico Medicine, Inc. A winner of multiple mathematical competitions and olympiads, he did his graduate work at the Mathematics & Mechanics Faculty of St Petersburg State Univ. E. Putin also works at the ITMO Univ. teaching Deep Learning courses and coaches Kaggle competition teams. He was one of the winners of the 6-day hackathon on deep learning, DeepHack.Game, at the Moscow Institute of Physics and Technology. E. Putin is one of the leaders of Phrama. AI, an Artificial Intelligence division of Insilico Medicine, where he is heading several projects evaluating aging biomarkers based on common blood biochemistry tests and transcriptomic data.

The availability of big data coupled with advances in highly-parallel high-performance computing led to a renaissance in artificial neural networks resulting in trained algorithms surpassing human performance in image and voice recognition, autonomous driving and many other tasks. While the adoption of deep learning in biomedicine has been reasonably slow, it is reasonable to expect major advances in personalized medicine, pharmaceutical R&D and other areas of healthcare. Here we present new results in developing aging biomarkers using blood biochemistry (www.Aging.AI) and transcriptomic data and demonstrate the applications of deep learning to drug discovery and drug repurposing using large data sets of transcriptional response data. We also discuss recent advances in deep learning and possible applications to actuarial science.

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

ISFA

STRUCTURING TEMPORAL SPARSE DATA WITH APPLICATION TO OPINION MINING

Julien VELCIN
ERIC, Lyon 2 University

Julien Velcin is associate professor of Computer Science at the University Lyon 2, with the qualification for supervising research (HDR, in French). He is a member of the Data Mining & Decision team of ERIC Lab. His current research focuses on machine learning with applications to the analysis of social media and opinion mining. He is particularly interested in weakly-supervised clustering techniques and topic modeling.

All the messages posted on the social media reflect only partially user's opinions. To gather those traces disseminated throughout the Web, evolutionary clustering techniques look eminently promising. In this talk, I will present two possible probabilistic models that address this issue. Our proposals extend the classic multinomial mixture for dealing with the temporal dimensions so that we can capture opinions over time. I will illustrate the two models with recent experiments performed within the ImagiWeb project that aims at studying the image (representation) of entities populating the social Web. I will more specifically use the image of two French politicians during the last presidential elections as a case study.

14:30

TREE-BASED CENSORED REGRESSION WITH APPLICATIONS TO INSURANCE

Xavier MILHAUD
ISFA, Lyon 1 University

15:20

Xavier Milhaud is currently associate lecturer at ISFA, Université Lyon 1. He was previously assistant professor at ENSAE ParisTech, partly in charge of the actuarial department. He did a PhD in AXA insurance company in which he mainly worked on understanding policyholders' behaviours in life insurance, more particularly focusing on modelling static and dynamic lapses. His research interests are closely linked to segmentation methods in order to model some portfolio heterogeneity by the use of regression models or non-parametric techniques. Typical practical applications are related to pricing and reserving in insurance.

In this paper, we propose a regression tree procedure to estimate the conditional distribution of a variable which is not directly observed due to censoring. The model that we consider is motivated by applications in insurance, including the analysis of guarantees that involve durations, and claim reserving. We derive consistency results for our procedure, and for the selection of an optimal subtree using a pruning strategy. These theoretical results are supported by a simulation study, and an application to insurance datasets.

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK   ISFA

ISFA-COLUMBIA
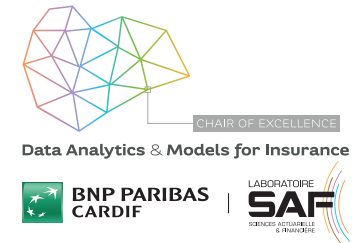
RESEARCH DISCUSSIONS WITH DAMI RESEARCH CHAIR TEAM

Data Analytics & Models for Insurance

CHAIR OF EXCELLENCE

BNP PARIBAS CARDIF | LABORATOIRE SAF

16:30

Christian ROBERT
ISFA, Lyon 1 University

Professor Christian Robert holds a PhD in applied mathematics from University Paris Denis Diderot, a MSc in Finance and Statistics, and is a former student of the Ecole Nationale de la Statistique et de l'Administration Economique (ENSAE). He is a fellow of the Institut des Actuaires. He is now full professor at ISFA, Université Lyon 1. Formerly, he was Associate Professor in Actuarial Science at ENSAE and Director of Graduate Studies at the Centre d'Etudes Actuarielles. Christian is an Associate Editor of the European Actuarial Journal. His main research interests include extreme value theory and statistics, actuarial theory and practice, and statistical finance.

Frédéric PLANCHET
ISFA, Lyon 1 University

Professor Frédéric Planchet holds a PhD in actuarial sciances from University of Lyon and is a former student of the École Nationale de la Statistique et de l'Administration Economique (ENSAE). He is a fellow of the Institut des Actuaires. He is now full professor at ISFA, Université Lyon 1. His main research interests include market consistent valuations, economic scenario generators (both risk-neutral and historical views) – Modelling censored and truncated data (mortality, disability, lapse, etc.): best estimate assumptions for reserve calculations, longevity risk evaluation, catastrophic risk in life insurance.

The "Data Analytics and Models for Insurance" chair of excellence has been created in september 2015 and aims at contributing to research in the fields of data analytics, risk management and governance for value creation in insurance companies. The chair works on the use of mathematical models for human decisions processes in insurance companies to enable them to adapt to a changing world, as well as the use of new flows of data on policyholders to model their behaviour and take it into account in forecasting models. This multidisciplinary chair involves researchers from different backgrounds (economists, actuaries, statisticians, data scientists and financial experts).

chaire-dami.fr

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK   ISFA

Feedback on Cardif Kaggle competition

Lam DANG
BNP Paribas Cardif

After graduating from AgroParisTech and ENSAE, Lam Dang started his carrer in 2013 as Data Analyst for RMS, where he did marketing and pricing analysis for numerous clients in retail industries. In 2015, he joined BNP Parisbas Cardif as Data Scientist, working on Machine Learning, Big Data and promoting Analytics. Lam is passionate for Data Science and has a big interest for Deep Learning and AI.

BNP Paribas Cardif launched its first challenge on Kaggle from Feb to April 2016.
The goal was to predict, early in the current claim process, the claims that BNP Paribas Cardif has appetite to accept much quicker, thus generating customer satisfaction.
The challenge was a success in term of participation (2926 teams), and solutions provided by top 3 winners.
In this presentation, Sébastien Conort (BNP Paribas Cardif's Chief Data Scientist) will talk about what was learned from winner's solutions.

16:30

RESEARCH DISCUSSIONS WITH LOLITA RESEARCH TEAM

Nicole EL KAROUI
Paris 6 University

Nicole El Karoui is currently emeritus professor of Applied Mathematics at the Laboratoire de Probabilités et Modèles Aléatoires of Pierre and Marie Curie University and previously professor at the École Polytechnique and Université du Maine (France). Her research has contributed to the application of probability and stochastic differential equations to modeling and risk management in financial markets. Nicole's research is focused on probability theory, stochastic control theory and mathematical finance and more recently on population dynamics and longevity risk.

Her contributions focused on the mathematical theory of stochastic control, backward stochastic differential equations (BSDEs) and their application in mathematical finance. In mathematical finance, she is known for her work on the robustness of the Black-Scholes hedging strategy, superhedging of contingent claims and the change of numéraire method for option pricing. Nicole is the coordinator of ANR Project LoLitA (Longevity with Lifestyle Adjustments) in Pierre and Marie Curie University.

Stephane LOISEL
ISFA, Lyon 1 University

Professor Stephane Loisel holds a PhD in applied mathematics from University of Lyon, a MSc in actuarial science and finance, and is a fellow and former member of the board of the Institut des Actuaires. He is now full professor at ISFA, University Lyon 1. He was visiting professor at ORIE, Cornell University in 2014 and has been lecturing for several years in Universite Paris 6 and ENSAE. Associate Editor of IME, MCA P, BFA , Risks and co-editor of EAJ, his main research interests include ruin theory with dependent risks, Solvency II, regulation and ERM, longevity risk and customer behaviour in insurance.

He is the coordinator of the ANR 4-year research project LoLitA (Longevity with Lifestyle Adjustments) and of the research chair Actuariat Durable sponsored by Milliman Paris. He received the SCOR PhD award in 2005, the Lloyd's Science of Risk runner-up prize in 2011 and the Hachemeister prize in 2013. Stephane also serves on the CERA review panel and is the scientific director of the French CERA program. He is a board member and member of the audit committee of April Group.

LoLitA

Olivier LOPEZ
Paris 6 University

Olivier Lopez is Professor of applied mathematics at Université Pierre et Marie Curie, Paris VI, Laboratoire de Statistique Théorique et Appliquée. His topics of research cover high-dimensional statistics, duration models, and copula theory.

The ANR project LoLitA aims at developing models for the uncertain long term development of human longevity and methods for managing longevity-related risk in pensions and long term health care. From a mathematical point of view, this requires advances in stochastic models for population dynamics and in certain classes of semi-Markov models, development of advanced numerical methods for such models, and development of new statistical methods (online change-point detection, calibration issues in longevity and long term care models,...).
The project is composed of six interconnected tasks, concerning respectively population dynamics modeling, long term care contracts, advanced simulation methods, multi-year solvency issues and joint stress tests, statistical aspects of longevity risk, and finally management of longevity risk in pensions.

lolita.isfa.fr

17:45

LATENT AND NETWORK MODELS WITH APPLICATIONS TO FINANCE

JingChen LIU
Columbia University

9:15

Jingchen Liu is Associate Professor at the Department of statistics at Columbia University. His research interests are Importance sampling, rare-event analysis; extremes of Gaussian random fields, queueing networks, and random ordinary and partial differential equations; latent variable modeling, cognitive assessment; computerized adaptive testing, sequential analysis; Bayesian modeling; missing data problems and multiple imputation; Markov chain Monte Carlo theory and applications.

One of the main tasks of statistical models is to characterize the dependence structures of multi-dimensional distributions. Latent variable model takes advantage of the fact that the dependence of a high dimensional random vector is often induced by just a few latent (unobserved) factors. In this talk, we present several problems regarding latent variable models. When the dimension grows higher and the dependence structure becomes more complicated, it is hardly possible to find a low dimensional parametric latent variable model that fits well. We further enrich the model by including a network structure on top of the latent structure. Thus, the main variation of the random vector remains governed by latent variables and the network captures the remainder dependence.

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

ISFA

APPLICATIONS OF DEEP LEARNING IN BIOMEDICINE, PART 2

Alexander ZHAVORONKOV
In Silico

See biography P.12

Evgeny PUTIN
Insilico Medicine

See biography P.12

The availability of big data coupled with advances in highly-parallel high-performance computing led to a renaissance in artificial neural networks resulting in trained algorithms surpassing human performance in image and voice recognition, autonomous driving and many other tasks. While the adoption of deep learning in biomedicine has been reasonably slow, it is reasonable to expect major advances in personalized medicine, pharmaceutical R&D and other areas of healthcare. Here we present new results in developing aging biomarkers using blood biochemistry (www.Aging.AI) and transcriptomic data and demonstrate the applications of deep learning to drug discovery and drug repurposing using large data sets of transcriptional response data. We also discuss recent advances in deep learning and possible applications to actuarial science.

10:30

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK   ISFA

DISTRIBUTIONAL ROBUSTNESS AND REGULARISATION IN MACHINE LEARNING: TWO SIDES OF THE SAME COIN

13:30

Karthyek MURTHY
Columbia University

Karthyek Murthy is a post-doctoral research scientist in the Department of Industrial Engineering & Operations Research at Columbia University. He completed his PhD at Tata Institute of Fundamental Research, Mumbai where he was the recepient of IBM International PhD fellowship and TCS Research fellowships. His research interests lie broadly in applied probability & stochastic processes, with special emphasis on models that arise in operations research, insurance and mathematical finance. Building on his PhD work on rare events,

he has been recently investigating stochastic modeling techniques that are robust to model risk. His PhD thesis was awarded with the TIFR-SASKEN Best thesis award for the year 2015

Finding the best fit given available data is a common theme encountered in various problems in machine learning. When the number of samples available for training is smaller than the ambient dimension of the problem, usual empirical risk minimisation may not be enough. We introduce RWPI (Robust Wasserstein Profile Based Inference), a novel machine learning methodology that is aimed at enhancing out-of-sample performance in such settings. RWPI exploits the relationship between a suitably defined distributionally robust optimization problem and the Wasserstein profile function. On one hand, a judicious choice of the distributional uncertainty can be used to build a wide range of regularisation procedures (we recover generalized Lasso, support vector machines and regularised logistic regression) as particular cases, and introduce new families. On the other hand, an asymptotic analysis of the Wasserstein profile function allows to optimally select the regularisation parameter. We shall discuss this optimality in the context of a popular regularised linear regression algorithm called generalized Lasso.

SRI INVESTMENT: APPLYING A SUPERVISED LEARNING ALGORITHM TO ENVIRONMENTAL, SOCIAL AND GOVERNANCE HISTORICAL SCORES

Christophe GEISSLER
Advestis & DAMI chair

Christophe graduated from «Ecole Normale Supérieure» Ulm (mathematics and computer science), and member of the French Actuaries Institute. He has occupied various positions in quantitative finance since 1986 in banks (Lazard, CPR, BNP Paribas, Société Générale) before launching Advestis in 2011. His fields of interest include quantitative finance, asset allocation models and machine learning. Christophe is also partner of Quinten, a french company created in 2008, applying data science to healthcare and insurance.

Socially Responsible Investment (SRI) relies on the inclusion of non-financial criteria, mainly environmental (E), social (S) and governance-related (G) scores attributed to corporations by specialized agencies. These scores are used to build a variety of quasi-static investment strategies like best-in-class or sector exclusion. The question of whether these strategies are associated with a positive or negative performance bias remains mostly opened.
We present here an attempt to apply supervised learning to a large set of co-variables comprising various derivations of initial score time series, in order to link future returns with observable variables. User requirements impose the traceability of the predictors, as well as their eligibility with respect to ESG fundamentals. An overview of the predictors provided by this approach, as well as the simulation of their systematic implementation is presented.

14:20

ON ROBUST RISK ANALYSIS

José BLANCHET
Columbia University

15:40

Jose Blanchet is a faculty member in the departments of IEOR and Statistics at Columbia University. Jose holds a Ph.D. in Management Science and Engineering from Stanford University. Prior to joining Columbia he was a faculty member in the Statistics Department at Harvard University. Jose is a recipient of the 2009 Best Publication Award given by the INFORMS Applied Probability Society and of the 2010 Erlang Prize. He also received a PECASE award given by NSF in 2010.

He worked as an analyst inProtego Financial Advisors, a leading investment bank in Mexico. He has research interests in applied probability and Monte Carlo methods. He serves in the editorial board of Advances in Applied Probability, Journal of Applied Probability, Mathematics of Operations Research, QUESTA, Stochastic Models, and Stochastic Systems.

We consider the problem of maximizing an expectation over all probability measures which are within a given Wasserstein distance of a target measure. This problem is solved explicitly for a large class of expectations of interest in great generality and we show that the solution has a natural practical interpretations in terms of stress testing. Moreover, when the underlying distribution is supported in R^{d} we show a weak limit for the asymptotic distribution of the empirical version of the problem. The limit laws are non-conventional and qualitative differences arise depending on the value of d (d=1, d=2, and d=3 give rise to three different mathematical developments). These results provide the foundation for a data-driven approach at robust stress testing. This talk is based on joint work with K. Murthy and Y. Kang.

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

ISFA

OVERREACTING RANDOM WALKS TEND TO BECOME AMNESIC

Nabil KAZI-TANI
ISFA, Lyon 1 University

Nabil Kazi-Tani holds a PhD in applied mathematics from Ecole Polytechnique in Paris, a MSc in mathematics and in economics and is a former student of Ecole Normale Supérieure in Cachan. He is now associate professor at ISFA, Université de Lyon 1.

Nabil specializes in probability theory, in the study of stochastic differential equations, dynamic risk measures and optimization, with applications in actuarial science and in particular in reinsurance models.

We are interested in a discrete random walk on integers that take the steps (1, +1) and (1, −1) with equal probability, when it is not equal to 0. When it reaches the state 0, the behavior of the walk depends on wether it touched the x-axis coming from a positive or a negative excursion.
The excursions of this walk do not form an i.i.d. sequence. We will give an explicit expression for the one dimensional distribution of this walk. As a by-product, we obtain a new simple combinatorial interpretation of k-fold convolutions of Catalan numbers.
We can for instance model an overreaction phenomenon at 0, meaning that the walk has a high probability to cross the x-axis when it reaches it. In that particular case, we will prove that in a financial model in discrete time where returns satisfy a symmetric overreaction property at 0, derivative prices are the same than in the model without overreaction.
We will also discuss the continuous time limit of such discrete random walks and explain how they are linked with the skew Brownian motion.
*This is a joint work with Dylan Possamaï (Paris Dauphine University).*
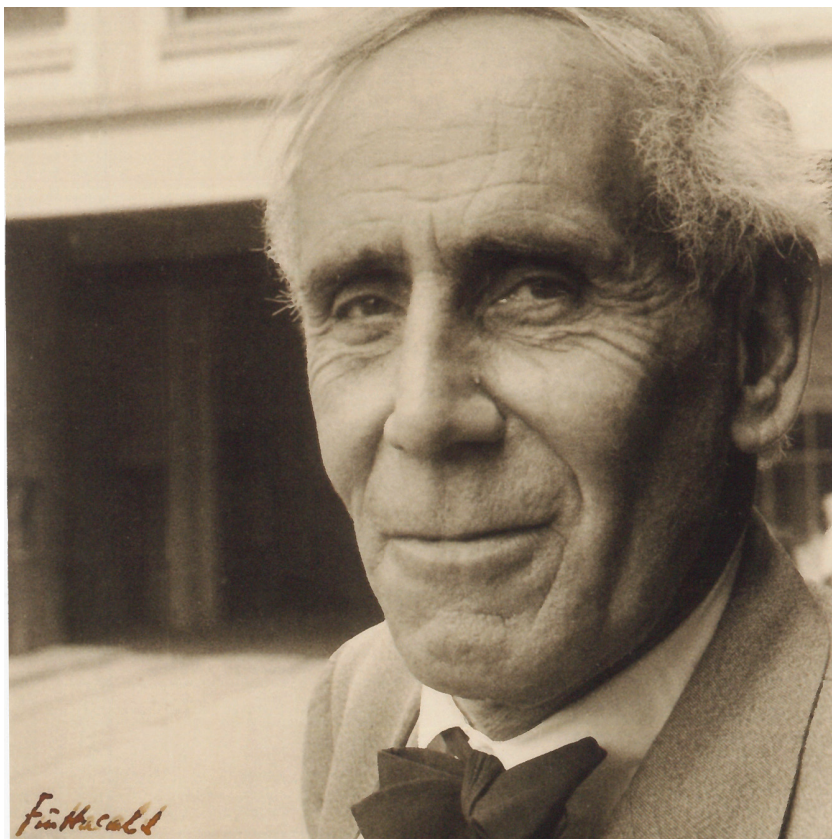
16:30

«I respect [Gumbel] even more as a person.
His political activity and his publications are sustained by a lofty ethos.
He is inspired by an uncompromising sense of justice.
Men like him are indispensable if we are ever to build
a sound political framework for our society».
ALBERT EINSTEIN

EMIL JULIUS GUMBEL

18 JULY 1891, MUNICH – 10 SEPTEMBER 1966, NEW YORK CITY

This photograph is courtesy of Patricia GUMBEL. Emil Julius GUMBEL (1891-1966) was a German mathematician and political writer. Following the murder of a friend, he attended the trial where he saw that the judge completely ignored evidence against the Brown Shirts Nazis. Horrified, he ardently investigated many similar political murders that had occurred and published his findings in Four Years of Political Murder in 1922. In 1928, he published Causes of Political Murder and also tried to create a political group to counter Nazism. GUMBEL was also one of the 33 signers of the 1932 Dringender Appell. Among the Nazi's most-hated public intellectuals, he was forced out of his position in Heidelberg in 1932. GUMBEL then moved to France, where he taught in Paris and at ISFA Lyon, and to the United States in 1940 where he taught at Columbia University in New York City until his death in 1966. As a mathematician, GUMBEL was instrumental in the development of extreme value theory, along with Leonard Tippett and Ronald Fisher. In 1958, GUMBEL published a key book on the topic: Statistics of Extremes. He derived and analyzed the probability distribution that is now known as the GUMBEL distribution in his honor.

EMIL J. GUMBEL
AN EXTREME STATISTICIAN

17:45

Matthias SCHERER
University of Munich

Prof. Scherer's (b. 1979) research area is mathematical finance and stochastics. The aim of his research is to appraise complex financial products and quantify their risks. His work mainly revolves around modeling dependency structures and assessing portfolio derivatives.

Prof. Scherer studied business mathematics at the University of Ulm. He obtained his Master of Science in mathematics at the University of Syracuse (USA).

He went on to do his doctorate in structural credit risk models at the University of Ulm (2007). In early 2007, he became coordinator of TUM's "Finance and Information Management" elite study program. Prior to his appointment as associate professor of mathematical finance (2010), he acted as interim professor for two semesters.

Gumbel's contributions to mathematical statistics can be roughly divided into three areas. At the beginning of his academic career, he mainly focused on descriptive statistics for population development and stochastic models for time of death distribution. This work, which was sometimes politically motivated (Gumbel 1923), has a significant political and historical value. He also demonstrated that many of the correlations obtained from craniometry are actually spurious. Gumbel also conducted studies into exceptionally long lives (Gumbel 1933). It is suspected that these analyses were what motivated him to devote his efforts increasingly to extreme value theory. In this second phase of his career – which lasted the longest and was the most ambitious academically – he worked on theoretical results about order statistics and the convergence of extreme observations, writing his most important works whilst in exile in France (Gumbel 1935). He was also involved in empirical work using data from hydrology and meteorology and popularised the methods of extreme value theory in various fields. His most creative period in extreme value theory came during his time in the USA. It culminated in a very well written and highly regarded series of lectures (Gumbel 1954) and the textbook «Statistics of Extremes» (Gumbel 1958). He continued his academic work until the age of 70 and, in the final years of his life, worked on bivariate probability distributions, motivated by linked models of survival probability. The copula that is named after him is taken from this work.

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK   ISFA

Important methods in quantitative risk management and actuarial science – particularly in the areas of extreme value theory and multivariate statistics – were developed and popularised by Emil J. Gumbel.

The Gumbel distribution and the Gumbel copula bear witness to this. To mark his 125th birthday – he was born on 18th July 1891 in Munich – we are providing an insight into his mathematical legacy and taking a closer look at his life as an academic, publicist, witness to history and pacifist.

In addition to his mathematical work, he published several books on politics and countless newspaper articles on political murders, the justice system and nationalist secret societies in the Weimar Republic.

This cost him his post at the University of Heidelberg in 1932 and saw him included on the first list of people whose citizenship was revoked under the Third Reich in 1933. He emigrated to France in 1932, but had to flee to the USA in 1940 to escape the occupying German forces. He died on 10th September 1966 in New York.

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

IEOR DEPARTMENT, COLUMBIA UNIVERSITY
340 W. Mudd Building
500 West 120 th Street, New York, NY 10027-6699

STATISTICS DEPARTMENT, COLUMBIA UNIVERSITY
1040 School of Social Work
1255 Amsterdam Av., New York, NY 10027
columbia.edu

ISFA

GRANDE ÉCOLE
D'ACTUARIAT
ET DE GESTION
DES RISQUES

INSTITUT DE SCIENCE FINANCIÈRE ET D'ASSURANCES

DOMAINE SCIENTIFIQUE DE GERLAND
50, Avenue Tony Garnier - 69366 Lyon cedex 7
Tél : +33 4 37 28 74 40 / Fax : +33 4 37 28 76 32
contact@isfa.fr - isfa.fr

Université Claude Bernard Lyon 1