Departement Statistiek en Aktuariële Wetenskap
Department of Statistics and Actuarial Science

Nuusbrief | Newsletter
July 2019

PAGE 14


PAGE 4


PAGE 5

IN THIS ISSUE

New Bachelor's degree in Data Science (BDatSci) from 2021

The field of Data Science is experiencing rapid and unplanned growth, spurred by the proliferation of complex and rich data in science, business and industry. Fuelled in part by reports such as the widely cited McKinsey report (2016) that forecasts a need for hundreds of thousands of Data Science jobs in the next decade, Data Science programmes have exploded in academia as university managements around the world have rushed to meet the demand. The website http://datascience.community/colleges currently lists 530 programmes in Data Science, Analytics and related fields at over 200 universities globally. The vast majority of these are master's degrees and certificate programmes offered both traditionally and online. There has also been a rapid growth of undergraduate programmes at both research institutions and major universities.

The Department of Statistics and Actuarial Science was instrumental in developing a 4-year undergraduate programme in Data Science. The programme has been tabled for internal approval at Stellenbosch University and we are currently in the process of submitting it for approval and accreditation by HEQC, CHE and SAQA. If successful, the first intake of students will be in 2021. In the meantime, students can register for BCom (Mathematical Sciences), with the focal area Data Science in 2020.

The goal of the BDatSci programme is to engage students in the full data cycle and to develop graduates with the necessary skills of a data scientist (Annual Review of Statistics, 2016).

This necessitates interweaving and integrating traditionally siloed topics and tools into a cohesive programme. We believe that at Stellenbosch University, a programme in Data Science at undergraduate level will definitely contribute and fulfil the need of an ever-changing world spurred on by the fourth industrial revolution. There is a real need for such an undergraduate programme in South Africa. This was further supported during a fact-finding tour of high-ranked American universities in October 2018 by the management team of Stellenbosch University. The key learnings report highlighted the need and importance of offering undergraduate and postgraduate programmes in Data Science in the USA and across the world. The job market demonstrated that the demand for data scientists is extremely high with South Africa already falling behind in this regard.

The new BDatSci programme is truly multi-disciplinary and will be offered in the following faculties: Science, Economic and Management Sciences, AgriSciences and Arts and Social Sciences. A total of eight focal areas have been identified within appropriate data-rich environments: Statistical Learning; Computer Science; Analytics and Optimisation; Applied Mathematics; Behavioural Economics; Statistical Genetics; Geoinformatics; and Statistical Physics.

A number of new modules in Data Science have been developed and will be introduced in each of the four years, while the fourth year will also include a Data Science research assignment of 40 credits. Students will use and apply the content of the different core and elective modules in the programme to conduct a research project under supervision of established researchers and specialists from industry. The exit level of this programme is NQF 8, which means that students will be able to register for a related master's degree in Data Science on completion.

For further information please contact Prof Paul J Mostert, pjmos@sun.ac.za

- *Annual Review of Statistics (2016). Curriculum Guidelines for Undergraduate Programs in Data Science, AA:1–26. {https://www.stat.berkeley.edu/~nolan/Papers/Data.Science.Guidelines.16.9.25.pdf}*
- *McKinsey Global Institute (2016). The age of analytics: Competing in a data-driven world. (In collaboration with McKinsey Analytics; www.mckinsey.com)*

ACADEMIC MATTERS

Industry involvement in academic programmes

Providing real-life context to the academic material covered in the various degree courses is important for the holistic development of our students' skills. One way of achieving this aim is facilitating engagement with lecturers with industry experience.

In 2019, the following actuaries have been or will be involved: Mr Peter Temple (GenRe, reinsurance), Mr Paul Lewis (Wider Fields), Ms Tanya van Niekerk (GenRe, reinsurance), Mr Karun Varughese (Sanlam, life insurance), Mr John Clur (GenRe, reinsurance), Mr Jean Gerin (Santam, short-term insurance), Mr Asher Grevler (Santam, short-term insurance), Ms Aliska van Niekerk (Dynamo, short-term insurance), Mr Hannes van Rensburg (Dynamo, short-term insurance) and Ms Lisa Schreiber (GRS Actuarial Consulting)

We are confident that our students are very well prepared for the fast-changing work environment.

Several of postgraduate courses in **Financial Risk Management** are also lectured by professionals from industry.

Ms Mary Roth, a Senior Quantitative Analyst at Old Mutual Specialised Finance, lectures a module in the Financial Risk Masters programme that covers hands-on problem solving using programming in Python and explores various applications thereof in Financial Risk Management. Mary obtained a MCom in Financial Risk Management at Stellenbosch University in March 2009.

Ms Mary Roth

Mr Francois Botha, Chief Investment Officer at Optimum Investment Group, lectures a module in the Financial Risk Masters programme which will broaden the students' knowledge of alternative investments. The course focuses on Hedge Funds, Private Equity, Structured Products and Real Assets. Students will be introduced to tools and techniques used to evaluate the risk-return attributes of each of the above-mentioned alternative investments. Francois obtained a MCom (Financial Risk Management) from Stellenbosch University in 2010.

Mr Francois Botha

Mrs Es-Marié Meyer (née Nortjé), a Decision Scientist (Quantitative Analyst) at DirectAxis, lectures the Value-At-Risk module in the Financial Risk master's programme. The subject matter studies various methods for the calculation of VaR, such as the parametric, historical and Monte Carlo approaches. Furthermore, it also looks at the calculation of VaR for option portfolios. Es-Marié graduated with a MCom in Financial Risk Management from Stellenbosch University in April 2014.

Mrs Es-Marié Meyer

Mr Henco Wiid, a Financial Risk Specialist / Consultant / Entrepreneur, lectures a module in the Financial Risk Masters programme which provides the students with an overview of credit risk from a scoring, accounting impairments and regulatory impairments perspective. The major topics covered in this module are: Introduction to Credit Risk Analytics SAS® Certified Statistical Business Analyst Using SAS® 9, Regression and Modelling - Online Course, Exploratory Data Analysis, Data Pre-processing for Credit Risk Modelling, Credit Scoring IFRS 9 in a nutshell, Probability of Default Loss Given Default and Basel 4 in a nutshell. Henco obtained a MSc in Quantitative Risk Management from North West University in 2012.

Mr Henco Wiid

Luca Steyn part of joint PhD programme with Ghent University

Luca Steyn, a lecturer and part-time doctoral student at Stellenbosch University, recently joined the joint PhD programme between Stellenbosch and Ghent University in Belgium. He tells us more:

"Postgraduate studies at Stellenbosch University hold many perks and one opportunity is the use of our exchange programme. As a result of previous collaboration between our Department and the Katholieke Universiteit, Leuven, I found a host that made it possible for me to visit Leuven for three months. During this visit I found a study leader for my doctoral studies at the Faculty of Bioscience Engineering, Ghent.

"I am based at the Department of Data Analysis and Mathematical Modelling in the Knowledge-based Systems (KERMIT) research group. This department focuses on data-driven and mathematical techniques to research biological phenomena. Different research fields include the BIOMATH, KERMIT and BIOBIX research groups. BIOMATH researches kinetic models, model reduction and model calibration with applications in resource recovery and sustainable pharmaceutical engineering. KERMIT focuses on knowledge-based, data-driven or spatially explicit modelling techniques. BIOBIX is the lab of Bioinformatics and Computational Genomics.

"My research here is related to open-set recognition with extreme value statistics. Open-set recognition is the ability to automatically detect when additional categories not known during the model-building phase are present in the data. Many biological application areas are open set by nature. One such area is the discovery of rare or alien species. The KERMIT research group collaborates closely with the Belgium research institute for nature and forest (INBO) who supplied a large dataset of animal species captured by trap cameras. The goal is to automatically detect new species that the current data does not contain.

"Being a Ghent student is an experience I value greatly. After regular working hours there are at least four hours left before sunset. This gives one plenty of time for social activities. Some of my favourite afternoon activities include rock climbing at the student gym, riding my bicycle through the wonderful city centre or having a famous Belgium beer with other doctoral exchange students.

"I would encourage all master's and doctoral students in the Department of Statistics and Actuarial Science to consider an exchange abroad. It gives one the opportunity to form lifelong connections with similar people and puts into perspective that Stellenbosch University is a world-class institution."

From left to right: Prof Dr Bernard de Baets (co-supervisor and head of the Department of Data Analysis and Mathematical Modelling, Ghent University), Mr Luca Steyn, and Prof Dr Stijn Luca (supervisor from Ghent University)

Successful first CARME conference in Africa

The Multivariate Data Analysis Group (MDAG) hosted the first Correspondence Analysis and Related Methods (CARME) conference outside of Europe in Stellenbosch, South Africa, from 4-6 February 2019. The three-day conference was attended by 68 delegates from 17 different countries. Plenary presentations were presented by Angela Montanari (University of Bologna), Peter Filzmoser (Vienna University of Technology), Riaan de Jongh (North West University), Frederic Lebaron (University Paris-Saclay), Cajo ter Braak (Wageningen University) and John Gower, via Skype (The Open University).

The scientific programme consisted of a wide variety of correspondence analysis applications: big data, compositional data analysis, CARME in finance, multivariate applications, CARME in the social space, principal component analysis, CARME in ecology, classification, biplots and bilinear models. Networking opportunities were also a highlight of the conference which started with a pre-conference excursion of the Cape Peninsula and a welcoming reception in the Botanical Gardens of Stellenbosch. The conference was concluded with a picturesque wine tasting outside of Stellenbosch and a gala dinner with traditional South African dishes.

Members of the local organising committee and speakers at the plenary presentations. First row: Johané Nienkemper-Swanepoel, Sugnet Lubbe, Michael Greenacre, Renette Blignaut and Niël le Roux. Second row: Peter Filzmoser, Stephan van der Westhuizen, Jörg Blasius, Riaan de Jongh and Cajo ter Braak.

Tiso Blackstar opens Business Day Premium subscription to Financial Risk Management students

In an effort to get our students to read the news and understand the theoretical contexts of their curriculums better, Dr Rousseau Lötter tried his luck and contacted Tiso Blackstar in the hope that they would open up some subscription services for his Risk Management classes. Tiso Blackstar obliged and opened up their Business Day Premium subscription for more than 170 second-year students, causing a groundswell of interest during a very interesting time for both international financial markets and local politics. Topical articles were debated in classes and even featured in the semester test. We would like to thank Tiso Blackstar for this generous sponsorship.

Financial Risk Management students

Graduation 2018/19: PhD and Master's Students

Prof Sarel Steel, Dr Francois Kamper and his father Prof Maarten Kamper

Student: Dr Francois Kamper
Degree: PhD (Mathematical Statistics)
Supervisors: Profs SJ Steel and JA du Preez
Title of thesis: Regularised Gaussian Belief Propagation

Abstract:

Belief propagation (BP) has been applied as an approximation tool in a variety of inference problems. BP does not necessarily converge in loopy graphs and, even if it does, is not guaranteed to provide exact inference. Even so, BP is useful in many applications due to its computational tractability. On a high level this dissertation is concerned with addressing the issues of BP when applied to loopy graphs. To address these issues, we formulate the principle of node regularisation on Markov graphs (MGs) within the context of BP. The main contribution of this dissertation is to provide mathematical and empirical evidence that the principle of node regularisation can achieve convergence and good inference quality when applied to a MG constructed from a Gaussian distribution in canonical parameterisation. There is a rich literature surrounding BP on Gaussian MGs (labelled Gaussian belief propagation or GaBP), and this is known to suffer from the same problems as general BP on graphs. GaBP is known to provide the correct marginal means if it converges (this is not guaranteed), but it does not provide the exact marginal precisions. We show that our regularised BP will, with sufficient tuning, always converge while maintaining the exact marginal means. This is true for a graph where nodes are allowed to have any number of variables. The selection of the degree of regularisation is addressed through the use of heuristics. Our variant of GaBP is tested empirically in a variety of settings. We show that our method outperforms other variants of GaBP available in the literature, both in terms of convergence speed and quality of inference. These improvements suggest that the principle of node regularisation in BP should be investigated in other inference problems. A by-product of GaBP is that it can be used to solve linear systems of equations; the same is true for our variant and in this context we make an empirical comparison with the conjugate gradient (CG) method.

Hyperlink to thesis http://hdl.handle.net/10019.1/104954

Student: Ms Chrismarie Enslin
Degree: MCom (Mathematical Statistics)
Supervisor: Prof SJ Steel
Title of thesis: Clustering methods with a focus on self-organising maps and an implementation on retail bank transactional data

Abstract:

The aims of this study are to provide an overview of traditional clustering methods, as well as introduce and discuss self-organising maps (SOMs) in detail. This study wants to convince the reader of the usefulness of self-organising maps as a dimension reduction tool. The batch SOMs algorithm was found to be the most appropriate SOM to use in practice, together with random initialisation of the prototypes. Ward linkage hierarchical clustering was found to perform the best on multivariate Gaussian simulated data and it was also found to be the most appropriate traditional clustering method to fit on top of the SOM. Banking transactional data was investigated for client behavioural clusters and the clusters of lower socio-economic class clients, technologically sophisticated clients, older and more traditional clients and low financial activity clients were found. These clusters emerged consistently throughout 9 different data samples.

STUDENT ACHIEVEMENTS

Student: Ms Loamie Kotzé
Degree: MCom (Mathematical Statistics)
Supervisor: Prof PJ Mostert
Title of thesis: Markov modelling of disease progression in the presence of missing covariates

Abstract:

The research investigates appropriate techniques for imputing missing data within the covariate structures when modelling the progression of breast cancer. A dataset was obtained from Isimo Health and based on this, data from the illness-death model was generated using Gumbel's bivariate exponential distribution with unit exponential margins. Various levels of missing at random within the covariate structure were implemented in these datasets to investigate imputation techniques. Two imputation techniques were applied on the simulated datasets, one based on chained equations (mice) and the other on random forests (missForest). Several performance measures were chosen to adequately identify the best imputation technique with a majority of the performance measures indicating that missForest imputation is superior. Based on these results, the random forest imputation technique was applied to the original dataset. Therefore, imputation based on random forests efficiently imputes missing data within the covariate structures of clinical data and can be applied when modelling disease progression.

Hyperlink to thesis http://hdl.handle.net/10019.1/106018

Student: Mr Jan Marais
Degree: MCom (Mathematical Statistics)
Supervisor: Dr S Bierman
Title of thesis: Deep learning for tabular data: An exploratory analysis

Abstract:

From about 2006, deep learning has proven to be very successful in application areas such as computer vision, natural language processing, speech and audio recognition, machine translation, bioinformatics, and social network filtering. These successes were undoubtedly facilitated by many advances in neural network architectures. In contrast, deep learning has not yet been found to excel in the context of tabular datasets. Many key machine learning tasks make use of tabular data, where currently the best machine learning models for tabular data use classification or regression trees as base learners. Therefore, the objective of this study is to identify, discuss and explore recent developments in deep learning which may be used to enhance the accuracy of deep neural networks in the tabular data domain. All major developments in the deep learning field are discussed and critically considered, with a view to improving deep learning in the context of tabular data. The challenges of applying deep learning to tabular data are identified, and on each of these fronts, potential improvements are proposed. The most promising modern deep learning architectures are further explored by means of empirical work. We also evaluate the validity of findings reported in the literature, and comment on the effectiveness of recent proposals. A useful by-product of the study is the development of a code base that may be used to implement the latest deep learning techniques, as well as for comparative model selection experiments.

Hyperlink to thesis http://hdl.handle.net/10019.1/106113

STUDENT ACHIEVEMENTS

Student: Ms Annegret Muller
Degree: MCom (Mathematical Statistics)
Supervisor: Prof SJ Steel
Title of thesis: Ensemble methods in multi-label classification

Abstract:

There are many scenarios where several labels may be associated simultaneously with each data case in a dataset. Therefore, a large number of multi-label datasets are found in a variety of domains, including image annotation, text annotation, bioacoustics, music research and medical diagnostics. In this thesis, we focus on such multi-label datasets with the particular goal of performing multi-label classification. Multi-label classification is an extension of binary- and multi-class classification to scenarios where each instance in a dataset can have multiple or none of K labels. Methods to perform multi-label classification can be divided into three categories: problem transformation methods, algorithm adaptation methods and multi-label ensemble methods. Multi-label ensemble methods receive particular attention in this research. We discuss previously proposed multi-label ensemble methods and propose a new multi-label ensemble method, named label dependent splitting (LDsplit) with trees. By performing an empirical study on benchmark datasets, the predictive performance of LDsplit with trees is compared to that of other multi-label learning methods. LDsplit with trees produces very promising results, allowing us to believe that with further modifications the procedure may become a highly competitive multi-label learning method. Furthermore, we also explore aspects of analysing text data. We perform an extensive analysis on a practical multi-label text dataset. The practical dataset consists of online comments, where each comment is labelled to identify if any or multiple so-called "toxicity" are present in the comment. Our model may therefore be used to identify different types of toxicity present in online comments and help reduce online abuse and harassment. A particular challenge faced in the practical data analysis is the sparsity of the labels.

Hyperlink to thesis http://hdl.handle.net/10019.1/104975

Student: Ms Jani Pretorius
Degree: MCom (Mathematical Statistics)
Supervisor: Prof SJ Steel
Title of thesis: Preconditioning for Feature Selection in Classification

Abstract:

Increased dimensionality of data is a clear trend that has been observed over the past few decades. However, analysing high-dimensional data in order to predict an outcome can be problematic. In certain cases, such as when analysing genomic data, a predictive model that is both interpretable and accurate is required. Many techniques focus on solving these two components simultaneously; however, when the data is high-dimensional and noisy, this approach may perform poorly. Preconditioning is a two-stage technique that aims to reduce the noise inherent in the training data before making final predictions. In doing so, it addresses the issues of interpretability and accuracy separately. The literature of this technique focuses on the regression case, but in this thesis, the technique is applied to a classification setting. An overview of the theory surrounding this method is provided, as well as empirical studies. Simulation studies evaluate the performance of the technique under various scenarios and compare the results to those obtained by standard (non-preconditioned) models. Thereafter, the models are applied to real-world datasets and their performances compared. Based on the results of the empirical work, it appears that, at their best, preconditioned classifiers can only reach a performance that is on par with standard classifiers. This contrasts with the regression case, where the literature has shown that preconditioning can outperform standard regression models in high-dimensional settings.

Hyperlink to thesis http://hdl.handle.net/10019.1/106057

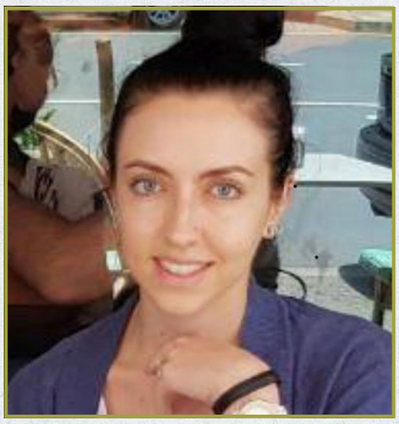
STUDENT ACHIEVEMENTS

Student: Mr Schalk van der Merwe
Degree: MCom (Mathematical Statistics)
Supervisor: Dr CJB Muller
Title of thesis: Application of statistics and machine learning in healthcare

Abstract:

Clinical performance and cost efficiency are key focus areas in the healthcare industry, since providing quality and affordable healthcare is a continuing challenge. The goal of this research is to use statistical analyses and modelling to improve efficiency in healthcare by focusing on readmissions. Patients readmitted to hospital can indicate poor clinical care and have immense cost implications. It is advantageous if readmissions can be kept to a minimum. Generally, stakeholders view strategies to address the clinical performance of healthcare providers, such as readmission rate, as mainly clinical in nature. However, this study investigates the potential role of machine learning in the improvement of clinical outcomes. This study defines machine learning as the identification of complex patterns (linear or non-linear) present in observed data, with the goal of predicting a certain outcome for new cases by mimicking the true underlying pattern in the population which led to the observed outcomes in the sample while throughout limiting rigid structural assumptions. The question at hand is whether patients that are at risk of readmission can be identified, along with the risk factors that can be associated with an increase in the likelihood of the event of readmission occurring. If yes, this can provide an opportunity to reduce the number of readmissions and thus avoid the resulting cost and clinical consequences. Once identified as a patient at risk for readmission, it will provide an opportunity for early clinical intervention. In addition, the model will provide the opportunity to calculate risk scores for patients, which in turn will enable risk adjustment of the readmissions rates reported. The data under consideration in this study is healthcare data generated by the operations of an international healthcare provider, Mediclinic International. The data that the research is based on is patient data captured on hospital level in all Mediclinic hospitals, operational in Mediclinic International's Southern African platform. The study aims to provide insight into machine learning techniques, as well as construct machine learning models that produce reasonable accuracy in terms of prediction of readmissions.

Hyperlink to thesis http://hdl.handle.net/10019.1/106084



Student: Ms Monique Beccaro
Degree: MCom (Financial Risk Management)
Supervisor: Prof WJ Conradie
Title of thesis: Overdraft Risk-Based Pricing and Profitability Model

Abstract:

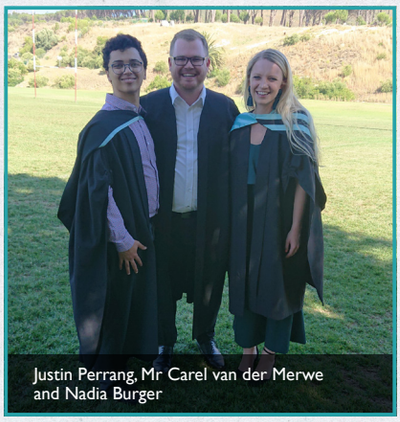
Risk-based pricing is the alignment of loan pricing with the expected loan risk in banks. A riskier customer should pay a higher price to compensate for the higher probability of default and cost to the lender. Pricing has a direct impact on revenue and capital adequacy and for this reason appropriate pricing is needed from a business and regulatory point of view. A loan application is accepted or rejected based on a customer's credit risk and this risk could also be used to determine the loan price. A risk-based pricing framework assists in identifying and managing the different risks within for example a Cheque Overdrafts portfolio of a bank. Such a framework will more accurately align the capital requirements with the actual amount of risk to which the bank is exposed. In this study the complete methodology implemented to complete a rebuilding process for the Cheque Overdrafts Pricing and Profitability model of one of the major banks in South Africa is discussed in detail.

STUDENT ACHIEVEMENTS

Student: Ms Monique-Mari Hugo
Degree: MCom (Financial Risk Management)
Supervisors: Prof WJ Conradie and Dr H Viljoen
Title of thesis: The Influence of Different Regimes on the Dynamic Estimation of GARCH Volatility Model Parameter Estimation

Abstract:

Volatility is used as the measure of risk within the financial markets. GARCH modelling involves important volatility forecasting methodology and is widely used in finance. In this study parameter estimates and volatility forecasts of GARCH models are investigated. Three GARCH models, the symmetric GARCH, GJR-GARCH and E-GARCH models, are compared. The dataset used in the study comprises of the JSE All-Share index. This index is divided into two different periods, namely, a tranquil financial period and a turbulent financial period. Different factors influence the performance of GARCH models and consequently determines which GARCH model is the most suited for certain circumstances. These factors are: the window period, forecasting horizon, the financial period, the underlying distribution of the log returns and the process of the mean equation. This study claims that a one-year window period is inadequate, and a five-year window period should be used. The 30-day forecasting horizon produces more accurate results than the one-day volatility forecasts. The normal distribution and the Student-t distribution are assumed for the daily log returns. Overall, the Student-t distribution produces marginally more accurate forecasts than the normal distribution.



Justin Perrang, Mr Carel van der Merwe and Nadia Burger

Student: Ms Nadia Burger
Degree: MCom (Financial Risk Management)
Supervisor: Mr N Dierick
Title of thesis: Stock Market Participation in the European Union

Abstract:

Understanding why most people shy away from stock market participation, both on individual and country level, is very important. Stock market participation explains consumption smoothing and this in turn can have a significant effect on household welfare. Literature has shown that welfare loss from withholding to invest in the stock market can be substantial. Furthermore, it has been estimated that limited stock market participation influences the distribution of wealth and wealth effects on consumption. A well-known phenomenon known as the stock market participation puzzle stems from the significance that most people do not invest in stocks despite the remarkable risk premium and gains from diversification involved. In this research, stock market participation both on household and country level are investigated for most countries in the European Union. This is done by using the Household Finance and Consumption Survey (HFCS). Influencing factors such as behavioural aspects and household characteristics are investigated. This adds to the growing literature on stock market participation especially in the field of behavioural and household finance. In this research a wide range of variables are considered in order to try and explain the stock market participation puzzle. This takes most prior literature, that only focuses on one determining factor, a step further.

Hyperlink to thesis http://hdl.handle.net/10019.1/104899

STUDENT ACHIEVEMENTS

Justin Perrang, Mr Carel van der Merwe and Nadia Burger

Student: Mr Justin Perrang
Degree: MCom (Financial Risk Management)
Supervisor: Mr CJ van der Merwe
Title of thesis: An application of copulas to improve PCA biplots for multivariate extremes

Abstract:

Principal Component Analysis (PCA) biplots is a valuable means of visualising high dimensional data. The application of PCA biplots over a wide variety of research areas containing multivariate data is well documented. However, the application of biplots to financial data is limited. This is partly due to PCA being an inadequate means of dimension reduction for multivariate data that is subject to extremes. This implies that its application to financial data is greatly diminished since extreme observations are common in financial data. Hence, the purpose of this research is to develop a method to accommodate PCA biplots for multivariate data containing extreme observations. This is achieved by fitting an elliptical copula to the data and deriving a correlation matrix from the copula parameters. The copula parameters are estimated from only extreme observations and as such the derived correlation matrices contain the dependencies of extreme observations. Finally, applying PCA to such an "extremal" correlation matrix more efficiently preserves the relationships underlying the extremes and a more refined PCA biplot can be constructed.

Hyperlink to thesis http://hdl.handle.net/10019.1/104861

Student: Mr Kiran Singh
Degree: MCom (Financial Risk Management)
Supervisor: Mr CJ van der Merwe
Title of thesis: Predicting Equity Movements using Structural Models of Debt Pricing and Statistical learning

Abstract:

Valuation is not an interesting problem in corporate finance, it is the only problem. Price and value are assumed to be the same number in economic theories of equilibrium and perfect capital markets. The economic theories of equilibrium asset pricing offer very weak practical suggestions for stock price behaviour at the firm level. The fundamental approach to stock price investing operates on the basis that price and value are two separate quantities and the stock price is fully determined by its intrinsic value. In this research the option-theoretic approach to default modelling is amended to provide an alternate view of value. Structural models apply an option-theoretic approach inspired by Merton (1974) that uses equity market and financial statement data in order to determine default probabilities. Default probabilities obtainable from the reduced form class of models provide the basis for extending the Merton model to estimate the firm value from market observable credit spreads. The probability of default is then a known constant provided from the reduced form model. The Merton model is reformulated with equity or firm value being used as the subject of the formula. The re-appropriated Merton model then provides a unique estimate of the firm's value based on current market information. The expected return on equity is then estimated from market credit spreads using individual capital structure and traded equity information. In this research it was found that historic estimates of return are poor predictors of future return at the firm level. The structural models provide good forecasts of return in some instances although have many challenges in implementation. The use of statistical learning methods was found to greatly improve predictions of future equity return movements using both debt and equity predictor variables, including unique predictor variables constructed from the structural models of the firm.

Hyperlink to thesis http://hdl.handle.net/10019.1/105931

STUDENT ACHIEVEMENTS

Student: Mr Jacob Kenyon
Degree: MCom (Statistics)
Supervisor: Dr D Hofmeyr
Title of thesis: Improving hyperplane-based density clustering solutions with applications in image processing

Abstract:

Minimum Density Hyperplane (MDH) clustering is a recently proposed method that seeks the location of an optimal low-density separator by directly minimising the integral of the empirical density function on the separating surface. This approach learns underlying clusters within the data in an efficient and scalable way using projection pursuit. The main limitation of MDH is that it defines clusters using a linear hyperplane. In recent research, MDH was applied to data which was non-linearly embedded in a high-dimensional feature space using Kernel Principal Component Analysis. While this method has shown to be an effective approach that extends the linear plane to a non-linear form, it does not scale well. A procedure is needed that can improve the hyperplane solution in an efficient way. We pose a novel approach to improve upon MDH by reassigning observations in a neighbourhood around a hyperplane solution using a gradient ascent procedure, Mean Shift. While Mean Shift is shown to provide promising results, the computation required to reassign objects becomes prohibitive as the size of the dataset increases. To reduce computation, a single step gradient heuristic is proposed whereby observations are reassigned based on the initial gradient evaluated at each point in relation to the hyperplane. This study critically reviews the validity of these approaches through applications with simulated and real-world datasets, with a focus on applications in image segmentation. We show that these approaches have the potential to improve hyperplane solutions.

Hyperlink to thesis http://hdl.handle.net/10019.1/106178

Student: Mr Agrippa Stulumani
Degree: MCom (Statistics)
Supervisor: Dr MMC Lamont
Title of thesis: Classification in high-dimensional data using sparse techniques

Abstract:

Traditional classification techniques fail in the analysis of high-dimensional data. In response, new classification techniques and accompanying theory have recently emerged. These techniques are natural extensions of linear discriminant analysis (LDA). The aim is to solve the statistical challenges that arise with high-dimensional data by utilising the sparse coding. In this project, our focus is on the following techniques: penalised LDA-L1, penalised LDA-FL, sparse discriminant analysis, sparse mixture discriminant analysis and sparse partial least squares. We evaluated the performance of these techniques in simulation studies and on two microarray gene expression datasets by comparing the test error rates and the number of features selected. In the simulation studies, we found that performance varies depending on the simulation set-up and on the classification technique used. The two microarray gene expression datasets are considered for practical implementation of these techniques. The results from the microarray gene expression datasets showed that these classification techniques achieve satisfactory accuracy.

Hyperlink to thesis http://hdl.handle.net/10019.1/105792

STUDENT ACHIEVEMENTS

Student: Mr William Mellville
Degree: MCom (Actuarial Science)
Supervisor: Mr SJ Burgess
Title of thesis: Investigation into the appropriateness of the South African microinsurance solvency capital regulation

Abstract:

The thesis provided a systematic comparison of the two different regulatory capital requirement options under the South African insurance regulatory framework, SAM, for insurers selling microinsurance. For the comparison, a model microinsurer was built that had been calibrated using industry data. It was found that the simplified microinsurance requirement resulted in less capital required than the requirement for traditional insurers. As would be expected, the simplified microinsurance requirement option was not as sensitive to changes in risk. However, using various deterministic and stochastic tests, both capital requirement options resulted in an insolvency probability of less than 0.5%, as required by SAM. The results also showed that it may, under certain circumstances, be acceptable to reduce the absolute minimum capital requirement for microinsurers. However, the reduction may not be sufficient to encourage informal funeral insurance providers to register with the regulator which was one of the goals of the microinsurance licence.

Hyperlink to thesis http://hdl.handle.net/10019.1/105817

Changes to the Actuarial Science curriculum

In 2018 we reported on changes to the Actuarial Science curriculum that are being phased in after a review of the international curriculum by the International Actuarial Association. Incorporating these changes also allowed us the opportunity to re-evaluate aspects of our actuarial offering.

This year changes were implemented to the second year of the BCom (Actuarial Science) degree. One of the interesting developments is the increased use of computer software, such as Excel, in problem solving. One of our second-year subjects will now also include a computer-based examination, where students will be required to use Excel to answer the questions in the examination paper.

In 2020 changes will be made to the third-year and honours Actuarial Science curriculums. An exciting development at the honours level is that more material will be included that is specifically relevant to actuaries working in South Africa. New material on Data Science will also be brought into third-year Mathematical Statistics.

Distinguished guests attend 2018 prize-giving ceremony

Every year in February the Department awards prizes to the best final-year student in each of the four subjects: Actuarial Science, Financial Risk Management, Mathematical Statistics and Statistics. The prize-giving ceremony coincides with the first meet-and-greet cocktail function for postgraduate students in the Department.

In 2018, our guest speaker was Mr Willem Roos, founder of Outsurance and current CEO of Rain, the new telecommunications company in SA. This year, we were very fortunate that Mr Gerrie Fourie, CEO of Capitec, and Mr Graham Lee, Head of Digital and Data Solutions at Capitec, accepted our invitation to address our students at the function. Over the years, we have treasured a close collaboration between Capitec and the Department. Various Data Science and Business Intelligence competitions were launched and co-hosted and many of our students are employed by Capitec. Gerrie and Graham gave interesting talks based on their own career paths and how one should start thinking about entrepreneurship. Students and staff thoroughly enjoyed their participation at this year's event.

What made this year's event even more special, is that the Rector and Vice-chancellor of Stellenbosch University, Prof Wim de Villiers, also accepted our invitation to attend the function. We were truly honoured by his presence at the event where Prof Ronel du Preez, Vice-dean: Teaching and Learning, handed over prizes to our four top students. Prof Niel Krige from the Alumni office at the University has also attended our functions for a number of years and we are humbled by his loyalty to the Department.

Prof Paul Mostert (Head of the Department), Mr Graham Lee (Capitec), Prof Ronel du Preez (Vice-dean: Teaching and Learning), Nicola Taylor (prize winner), Danielle Vellema (prize winner), Prof Wim de Villiers (Rector of Stellenbosch University), Mr Gerrie Fourie (CEO of Capitec) and Cameron Hendriks (prize winner).

The 2018 prize winners are:

Statistics – Nicola Taylor

Nicola grew up in Johannesburg and matriculated from Crawford College. She enjoys doing fun runs, going to wine farms with friends and participating in the Investments Society, among other things. Nicola is currently enrolled for an Honours degree in Statistics.

Why did you decide to study Statistics?

The field of Statistics is so broad and interesting, and I enjoy the challenging nature of the work. It is a field that is forever changing and has so many applications. I find this fascinating and think that the skills we learn in Statistics at university can be used in almost any business.

Do you think Statistics is a useful subject that can bring change to our society at large?

Yes, there are so many incredible directions to which the field of Statistics can be applied. We live in a world of information and data and need people and businesses who know how to process and interpret it. This can assist communities, companies and governments to function more efficiently, directing resources to where they are most needed and at a lower cost. For example, Statistics is crucial in the field of medicine: many are reliant on prescribed drugs for cancer or HIV, and statisticians interpret clinical data to ultimately improve chronic illness allocation of resources and achieve better outcomes.

STUDENT ACHIEVEMENTS

What are your career plans?

I will build on my undergraduate degree in Investment Management and advance my career in Statistics abroad in 2020.

Financial Risk Management – Daniel de Kock

Daniel was born in Cape Town, but moved to the island of Guernsey, a Crown dependency in the English Channel, when he was eight. He attended Elizabeth College in Guernsey, which was founded in 1563 by Queen Elizabeth I, and is one of the oldest private schools in the British Isles. He is currently working in Finance in Guernsey.

Why did you decide to study Financial Risk Management?

As Guernsey is classified as 'offshore', it is a large financial hub, thus I have been immersed in Finance from a young age. This combined with an affinity and love for mathematics made Financial Risk Management my perfect focal area.

What are your career plans?

I am currently busy with my Chartered Institute For Securities & Investment (CISI) – Investment Advice Diploma, which once completed would make me RDR Qualified. Within the next few years I would like to become a CISI Chartered Wealth Manager. I am very pro-Africa when it comes to investment and I continually follow African news and business opportunities, and eventually I would like to 'head' a team focusing on Africa.

What are some of your hobbies, other interests and achievements?

After leaving school I spent one and a half years playing rugby at provincial and championship level, receiving honours with Western Province U19, Leicester Tigers U20 and Doncaster Knights. During my time at Stellenbosch I played for various teams: Maties, Victorians, Cravens and Markotters. In recent years I also took part in competitive powerlifting, qualifying for the South African National Championships in April 2018, and briefly holding the U90 kg Junior South African deadlift record of 260 kg. I am currently still actively training in both powerlifting and rugby.

Actuarial Science: Danielle Vellema

Danielle grew up in George and matriculated from Outeniqua High School. One of her favourite things is to escape the rush and frenzy by immersing herself in nature. She loves the Kruger National Park and scuba diving. Her hobbies also include music and ballet. Living in Stellenbosch for the past three years, she considers herself a developing 'foodie' and budding oenophile. She is currently enrolled for an Honours degree in Actuarial Science.

Why did you decide to study Actuarial Science?

I decided to study Actuarial Science because of the challenging nature of the course, as it has become notorious as arguably the most demanding degree at university. As an internationally recognised qualification it also offers many career opportunities.

STUDENT ACHIEVEMENTS

Do you think Actuarial Science is a useful subject that can bring change to our society at large?

Most definitely. We are living in increasingly uncertain times and I believe actuaries are especially equipped to make a difference in providing some security to society as a whole; for example, it allows for much more accurate projections regarding future challenges for our developing democracy to enable improved planning regarding the physical, mental and social well-being of an aging population.

What are your career plans?

I am very interested in the healthcare field of Actuarial Science, especially within the South African context. We have such an incredible challenge in improving access to quality medical care for all; we need more innovative funding strategies and I believe actuaries will be able to provide invaluable contributions to the future success of universal health care in SA.

What do you like most about the Department and your postgraduate programme?

The best thing for me is the incredibly talented lecturers that we now get to interact with on a much closer level than as undergraduates.

Mathematical Statistics: Cameron Hendriks

Cameron was born in George but moved to Hermanus when he was five years old and matriculated from Hermanus High School. He is a member of the rowing executive for Maties Rowing and is an avid weightlifter. He regularly partakes in CrossFit competitions and enjoys testing his physical skills. Cameron is currently enrolled for an Honours degree in Actuarial Science.

Why did you decide to study Mathematical Statistics?

I have always had a passion for solving complex problems and applying anything I learnt in school to the real world. Mathematics gave me a tool set in doing so, but Mathematical Statistics taught me how to get the most value from them.

Do you think Mathematical Statistics is a useful subject that can bring change to our society at large?

Of course! Many of the modern problems we face today stem from uncertainty. I believe that many of these problems could be solved or further insight could be gained if statisticians collaborate with professionals in various fields. I believe that the true value of Statistics stems from its contribution to solving pressing societal problems rather than the complexity of its tools.

What are your career plans?

I intend on completing my Honours degree in Actuarial Science, after which I plan on pursuing a career in investments. I am intrigued by private equity and the uncertainty surrounding it. I hope to be able to add value in this sector.

What do you like most about the Department and your postgraduate programme?

I consider the Department of Statistics and Actuarial Science at Stellenbosch University to be the best in the country. Now at a postgraduate level, after the technical foundation has been laid in undergrad, I am excited to apply my skills in interesting ways, such as simulation and financial economics to name a few.

2019 CGW Schumann medal recipient: Niel Botha

The prestigious CGW Schumann medal is awarded annually to the best postgraduate student in the Faculty of Economic and Management Sciences. For a second year in a row, this medal was awarded to a student from the Department of Statistics and Actuarial Science.

Mr Niel Botha completed the degree BCom (Actuarial Science) (cum laude) in 2016, passing all his subjects with distinction. In 2018, he obtained a BComHons (Actuarial Science) degree, the only student in this class to pass the degree cum laude. Based on the subjects he has completed to date he will be recommended for all the available exemptions from the examinations of the Actuarial Society of South Africa based on performance in equivalent university modules taken in his undergraduate and honours degrees. Niel recently also passed two of the final (Fellowship) actuarial examinations written through the Actuarial Society of South Africa and has now only the final Fellowship examination to pass in order to complete all the examinations required to become an actuary.

What makes his academic achievement even more exceptional is that Niel played cricket at a very high level throughout his university career, representing Maties as well as playing at provincial level for Boland and the Cape Cobras. This shows how well Niel is able to manage his time and arrange his priorities, as such achievements do not come without a great deal of effort.

Niel is also a natural leader, leading by example and encouraging others to succeed. He has captained numerous sports teams and has always been willing to take the lead in class work. He served on his high school's student council, and while at university he served on the Golden Key sub-committee involved with the planning and implementation of a tutoring programme for disadvantaged scholars in Kayamandi. Niel was involved in various outreach programmes, providing assistance to those in less fortunate circumstances. Such programmes have included the tutoring of mathematics at Makupula High School in Kayamandi, providing food and clothing to the homeless, and providing cricket coaching and mentoring to local underprivileged children.

Niel will be heading to Cambridge University (St. Catherine's College) in September to study towards a master's degree. He has been awarded the prestigious Skye Foundation Scholarship to fund his studies. The Department is very proud of Niel and wish him all the best with his new endeavours in the United Kingdom.

Prof Garrett Slattery (head of Actuarial Science), Prof Ingrid Woolard (Dean of the Faculty of Economic and Management Science), Mr Niel Botha (medal recipient), Prof Ronel du Preez (Vice-dean Teaching) and Prof Paul Mostert (Chairperson of the Department)

More students benefit from Schroders and Correlation Risk Partners bursaries

Bursaries funded by Schroders and Correlation Risk Partners were awarded to 21 final-year and postgraduate students in the Department of Statistics and Actuarial Science at Stellenbosch University (SU) in 2019.

This brings the total of bursaries awarded to Stellenbosch students over the past four years to 67. An agreement between SU's Department of Statistics and Actuarial Science and Schroders and Correlation Risk Partners, which comprises postgraduate bursaries and research collaboration, was signed in November 2015. This year's bursaries were awarded at a special dinner on 10 June.

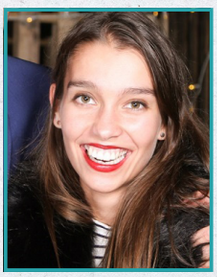
Schroders bursary holders with officials from Schroders, Correlation Risk Partners and Stellenbosch University.

NRF-SASA Academic Statistics Bursaries

In 2014, the national Department of Science and Technology (DST) classified Statistics as a "vulnerable discipline" in South Africa due to the scarcity of well-qualified professionals and academics in this field. Under the NRF division of Knowledge Advancement and Support (KAS), a grant was specifically allocated to address the crisis in academic Statistics, with an overall goal of attracting bright young minds back to the academic profession.

Administered by the NRF and the South African Statistical Association, the fund is used to provide full-time bursaries to master's and PhD students. The only criterion to obtain a bursary is academic merit. Stellenbosch University's third master's cohort (2017/2018) graduated at the end of 2018 and in March 2019, while the third cohort of PhD students will graduate at the end of 2019. The Department has appointed three beneficiaries under this funding initiative as permanent lecturers, namely: Dr Trudie Sandrock, Dr Francois Kamper and Mr Luca Steyn.

The recipients of the prestigious NRF bursaries in 2019 are Sven Buitendag and Annegret Muller (both PhD students), as well as Bronwyn Dumbleton, Michail Melonas and Corine de Koker (all master's students).


Recipients of the Faantjie and Lettie Pretorius bursary for undergraduate students

The heirs of Prof Faantjie Pretorius, the first professor of Statistics at Stellenbosch University and founder of the Department of Statistics and Actuarial Science, donated a substantial amount of money to the University. The return of this investment is used to award bursaries to deserving students in the Department every year. In 2019, Jana Nieuwenhuizen and Freddie van der Merwe share the bursary from this fund.

Recipients of the Novare bursaries for Financial Risk Management students

The Department of Statistics and Actuarial Science has maintained a very good professional working relationship with Novare since the independent investment advisory business was founded in South Africa in 2000. It has since expanded into Africa. Mr Derrick Roper, a qualified actuary and alumnus of the Department is one of the founding directors of Novare. Through the years, Novare has supported the Financial Risk Management programme substantially in the form of significant financial contributions. This year's donation was used to award bursaries to seven undergraduate students in Financial Risk Management. They are from left to right: Samuel Sendzul, Elizabeth Rajah, Elrencia Carinus, Wilmari van Dyk, Cara Weyers, Melissa Carstensen and Daylan Leibrandt.

STUDENT BURSARIES

Every year, the Department awards bursaries to deserving postgraduate students. Three bursaries were awarded this year. Although the bursaries aren't full academic bursaries, they will supplement the financial needs of students. The recipients are: Ingo von Gossler, Ishmat Ahmed and Thamusanqa Mnyulwa. All three are Honours students in Statistics.

Ingo von Gossler, Ishmat Ahmed and Thamu Mnyulwa

Stellenbosch alumni become Fellow members of the Actuarial Society

Being admitted as a Fellow of the Actuarial Society of South Africa (FASSA) is a big achievement, which requires many years of dedicated study. The following alumni from Stellenbosch University have been admitted as Fellows and the Department would like to congratulate them:

Surname	Initials	Surname	Initials
Appelo	H	Nicholas	NN
Brand	IJ	Oakley	CAS
Carse	DF	Scannell	Ina-Mari
Crous	JS	Schoeman	JM
du Plessis	C	Strydom	S
Klue	PE	van der Berg	RS
MacPherson	C	van der Merwe	JMJ

Meet Sandra Loggenberg

She ensures that all our offices are clean, our kitchen and staff room are in order and greets everyone with a friendly smile. Sandra Loggenberg has been one of the Department's assistants for the past 13 years.

"I am very happy in my work and feel at home here. The staff in the Department care a lot about me," she says.

She was born and raised in Idas Valley where she still lives. Sandra is the mother of three children (one of whom is Denver Loggenberg, another assistant in the Department) and grandmother of two grandchildren. According to Sandra she loves cleaning and making sure everything is neat, whether it is at home or at work.

"I like creating a home environment where everything is in order. I want to be seen as a person who is neat, friendly and caring. These are the things I like about people."

Prior to being a departmental assistant, Sandra was an assistant at Majuba men's residence for many years where she prepared food for students and made sure that their rooms were clean. Sandra enjoys travelling with her sister and together they have travelled to many places in South Africa.

"I still hope to travel a lot more in the future," she adds.

Tertius de Wet talks about Stellenbosch Feeding in Action

Prof Tertius de Wet is involved with the organisation called Stellenbosch Feeding in Action as Chairman of the Board of Directors. In this article, Tertius tell us about the important work this organisation has been doing for many years in the Stellenbosch area.

Stellenbosch Feeding in Action ("Stellenbosch Voedingsaksie") was established in 1994 as a programme to help poor and vulnerable people in the Stellenbosch area meet their most basic need: to have sufficient and nutritious food. The founder, Dr Dora Steyn, realised there was a great need for such support in the local community and started with a single outreach point at Bergzicht taxi rank. It grew over the years with the aim to provide nutritious and satisfying food to anyone who is hungry, whether they are members of struggling households or people who are homeless, sick or vulnerable in the belief that no hungry person should ever be turned away. The focus is especially on feeding vulnerable children who need sufficient food and nutrition to grow and stay healthy by providing meals at schools, crèches and aftercare centres in the Stellenbosch area.

Currently Feeding in Action supplies approximately 5 000 people per week with meals at 45 feeding points covering the broader Stellenbosch municipal area from Franschhoek to Jamestown. This works out at 10 000+ meals per week at these feeding points. Furthermore, 1 000+ food parcels are supplied per week at clinics, schools and other organisations. A number of projects are undertaken, such as food support to different holiday projects organised for learners.

Learners from local primary and high schools often carry out their community projects with Feeding in Action, preparing soup and sandwiches and assisting with the distribution and serving of meals.

The staff of Feeding in Action is supported by approximately 120 volunteers from the communities they serve, assisting with the processing of vegetables, preparing of meals and serving of meals at the different feeding points. Apart from the community volunteers, approximately 200 volunteers from Stellenbosch assist in all aspects of the organisation's activities. For further information visit www.feedinginaction.co.za

Congratulations are in order

In the words of the novelist Dostoyevsky "the soul is healed by being with children" and three staff members within the Department have recently received an additional dose of healing this year.

Early in January, departmental assistant Denver Loggenberg became a proud father (for the first time) to a baby daughter, Amia. The new family (together with mom, Aneshia) are doing well. However, on randomly occurring mornings Denver's usual high energy level appears to be just a little subdued, sleep being a prized (and sometimes elusive) commodity.

Another staff member, Rousseau Lötter, a senior lecturer in Financial Risk Management, was blessed with the birth of his second daughter. Cara Bella Lötter was welcomed into the world by her parents on 30 January. Rousseau, always beaming with pride, has taken his new role of father to two wonderful daughters in his stride.

Chris Muller and his wife Liesl welcomed a new addition to their family in June 2019. Their daughter, Chrisie Jane Beatrice Muller, became the most recent in a long line of Mullers to share the initials "CJB". Even though Chris is beaming with pride, with two young daughters to dote on he will probably only find time for himself in bits and bytes in the near future.

As a department, we would like to congratulate our staff members on the new additions and wish them lasting joy.

UPCOMING EVENTS

Bayesian Biostatistics workshop: 18-22 November 2019

Prof Emmanuel Lesaffre of the Leuven Biostatistics and Statistical Bioinformatics Centre (L-BioStat), Katholieke Universiteit Leuven, Belgium, will present a five-day workshop course on Bayesian Biostatistics. The workshop is partly funded by the recently established Universities Capacities Development Programme (UCDP): National collaborative project – Strengthening academic staff development in Mathematical and Statistical Sciences in South Africa. The workshop will be held in one of the postgraduate venues of the Department of Statistics and Actuarial Science.

Prof Lesaffre is no stranger to South Africa and particularly Stellenbosch, having presented a number of workshops on Bayesian Biostatistics and Longitudinal data analysis. He will also present a one-day workshop during the annual South African Statistical Association conference at Nelson Mandela University in November.

UCDP beneficiaries from all participating Universities can attend this year's workshop at no cost. The UCDP through DHET will fully sponsor the registration cost, travel and accommodation costs covering the five days. Academics who qualify under the UCDG should be permanently employed at a Statistics Department, busy with a master's or PhD degree in Statistics, or should have obtained a PhD in Statistics within the last five years. Academics should also be on the UCDG database that was updated by the node leaders of each participating University.

Registration cost: R4 000 including all notes, coffee/tea and lunch every day. Wi-Fi connection available to all delegates for the duration of the workshop.

Delegates should arrange their accommodation and transport through own institution and claim for all expenses through the respective node leader if using the UCDP funds.

Information about the workshop will soon be updated on the official website of the workshop. For further enquiries contact: Dr Chris Muller at cmuller@sun.ac.za

SUMMARY OF WORKSHOP:

For many decades, the frequentist approach dominated statistical practice. Despite that the Bayesian approach saw the light more than 250 years ago, it became only (and gradually) attractive for the statistical practitioner the last ten to twenty years. The approach is especially known for its capability to tackle complex statistical modelling tasks. However, the Bayesian approach is nowadays also recognised as an alternative, and sometimes better tool, than the frequentist approach in experimental (clinical) research. In this five-day course the Bayesian ideas and tools will be explained. The basics of Bayesian methodology will be explained on the first two days of the course. The next three days are devoted to a great variety of Bayesian methods in primarily medical applications covering: diagnostic tests, clinical trials, longitudinal studies, (network) meta-analyses, survival analysis. The course is scheduled into classroom teaching and computer exercises, and uses the software packages Win/OpenBUGS and JAGS, possibly in combination with R. The course is partly based on the Wiley book published in 2012 of Lesaffre and Lawson, entitled Bayesian Biostatistics.

The course assumes a good knowledge of regression techniques (linear, logistic, etc.) and some knowledge of models for correlated data. Programming skills and experience with R are beneficial though not essential for the course.

Emmanuel's research interests include Bayesian methods, longitudinal data analysis, statistical modelling, analysis of dental data, interval censored data, misclassification issues and clinical trials. He has written more than 500 papers in peer-reviewed statistical and medical journals. He is the founding chair of the Statistical Modelling Society, past president of the International Society for Clinical Biostatistics and fellow of ISI and ASA. He (co-)authored seven books among which Bayesian Biostatistics (2012) with Andrew Lawson and the recently published Chapman and Hall/CRC book Survival Analysis with Interval-Censored Data: A Practical Approach with examples in R, SAS, and BUGS (2017), together with Kris Bogaerts and Arnost Komárek.

UPCOMING EVENTS

Upcoming seminars

SEMINAR PROGRAM: SECOND SEMESTER 2019	
Stellenbosch University Department of Statistics and Actuarial Science	
26 July	Humphrey Brydon (Department of Statistics and Population Studies, UWC) Exploration of missing data imputation methods
16 August	Stefan Britz (Statistical Sciences, UCT) Computer vision applications: from Cedar tree counting to precision farming, GTA simulations and going ballistic(s)
30 August	Rulof Burger (Department of Economics, SU) An Econometric method for estimating population parameters from non-random samples: an application to clinical case finding
27 September	Daan le Roux (Department of Information Science, SU) Technology use effects: current evidence and challenges
4 October	Ian Louw (Explore Data Science Academy, Cape Town) Water network management: a data-driven approach
18 October	Martin Coxon (Department of Statistics and Actuarial Science, SU) Identifying the utilisation trends of diabetic beneficiaries in the South African private healthcare sector

Lectures start at 13:00 and are held in room 2048 of the Van der Sterr Building, c/o Victoria and Bosman streets, Stellenbosch.

Enquires: Prof Danie Uys
Telephone: 021 808 3879
E-mail: dwu@sun.ac.za

Data science and machine learning reading group

The Data Science and Machine Learning reading group in the Department is a group of researchers whose interests intersect in the challenges associated with the treatment and handling of modern data-based problems. The group has broad interests in the statistical theory and methodology associated with such problems, including dimension reduction and visualisation; probabilistic graphical models; classification; clustering and model selection, as well as diverse areas of application, including wide-scale spatial mapping of soil properties; music information retrieval; image analysis and segmentation; pneumonia progression in infants; tumour growth prediction based on cancer cell-lines and lifetimes of table grapes. For more information about our research and activities, please contact Dr David Hofmeyr (dhofmeyr@sun.ac.za).

List of published articles and presentations

Published papers in peer-reviewed journals in 2018

- Claassen-Weitz, S., Gardner-Lubbe, S., Nicol, P., Botha, G., Mounaud, S., Shankar, J., Nierman, W.C., Mulder, N., Budree, S., Zar, H.J. and Nicol, M.P., 2018. HIV-exposure, early life feeding practices and delivery mode impacts on faecal bacterial profiles in a South African birth cohort. Scientific reports, 8(1), p.5078.

- Gower, J.C., Gardner-Lubbe, S. and Le Roux, N.J., 2018. Interaction: Fisher's Optimal Scores Revisited. Journal of Agricultural, Biological and Environmental Statistics, 23(1), pp.92-112.

- Pretorius, E.J., Zarrabi, A.D., Griffith-Richards, S., Harvey, J., Ackermann, H.M., Meintjes, C.M., Cilliers, W.G., Zunza, M., Szpytko, A.J. and Pitcher, R.D., 2018. Inter-rater reliability in the radiological classification of renal injuries. World journal of urology, 36(3), pp.489-496.

- Innes, S., Harvey, J., Collins, I.J., Cotton, M.F. and Judd, A., 2018. Lipoatrophy/lipohypertrophy outcomes after antiretroviral therapy switch in children in the UK/Ireland. PloS one, 13(4), pp.1-10.

- Kamper, F., Du Preez, J.A., Steel, S.J. and Wagner, S., 2018. Regularized Gaussian belief propagation. Statistics and Computing, 28(3), pp.653-672.

- Koen, N., Philips, L., Potgieter, S., Smit, Y., Van Niekerk, E., Nel, D.G. and Visser, J., 2018. Staff and student health and wellness at the Faculty of Medicine and Health Sciences, Stellenbosch University: current status and needs assessment. South African Family Practice, 60(3), pp.84-90.

- Koen, N., Wentzel‐Viljoen, E., Nel, D. and Blaauw, R., 2018. Consumer knowledge and use of food and nutrition labelling in South Africa: A cross‐sectional descriptive study. International journal of consumer studies, 42(3), pp.335-346.

- Leclezio, L., Gardner-Lubbe, S. and de Vries, P.J., 2018. Is It Feasible to Identify Natural Clusters of TSC-Associated Neuropsychiatric Disorders (TAND)?. Pediatric neurology, 81, pp.38-44.

- Lötter, R., 2018. Recommendation pattern preferences and share analysts bias on the JSE. Studies in Economics and Econometrics, 42(1), pp.1-22.

- Marx, M., Young, S.Y., Harvey, J., Rosenstein, D. and Seedat, S., 2017. An examination of differences in psychological resilience between social anxiety disorder and posttraumatic stress disorder in the context of early childhood trauma. Frontiers in psychology, 8, p.2058.

- Pepler, P.T., Uys, D.W. and Nel, D.G., 2018. Regularized covariance matrix estimation under the common principal components model. Communications in Statistics-Simulation and Computation, 47(3), pp.631-643.

- Rossouw, R.F., Coetzer, R.L.J. and le Roux, N.J., 2018. Simulation of a coal stacking process using an online X-Ray Fluorescence analyser. ORiON, 34(1), pp.65-81.

- Rudman, M., Leslie, A.J., van der Rijst, M. and Hoffman, L.C., 2018. Quality characteristics of Warthog (Phacochoerus africanus) meat. Meat science, 145, pp.266-272.

- Van der Merwe, C.J., Heyman, D. and de Wet, T., 2018. Approximating risk-free curves in sparse data environments. Finance Research Letters, 26, pp.112-118.

- Williams, R., van Heerden, J.D. and Conradie, W.J., 2018. Value at risk and extreme value theory: application to the Johannesburg Securities Exchange. Studies in Economics and Econometrics, 42(1), pp.87-114.
- Minkah, R., de Wet, T. and Kwabena, D.A., 2018. On Extreme Value Index Estimation under Random Censoring. African Journal of Applied Statistics, 5(2), pp.419-445.
- Minkah, R., de Wet, T. and Nortey, E.N.N., 2018. A Simulation Comparison of Estimators of Conditional Extreme Value Index under Right Random Censoring. African Journal of Applied Statistics, 5(1), pp.337-359.

Academic papers presented in 2018/19

Presentations at local conferences

SOUTH AFRICAN STATISTICAL ASSOCIATION (SASA) CONFERENCE, 26-29 NOVEMBER 2018, JOHANNESBURG:

- Nienkemper-Swanepoel, J., le Roux, N.J. and Lubbe, S. Visualisations for identifying missing data mechanisms. Annual South African Statistical Association Conference, 26-29 November 2018, Johannesburg.
- Sandrock, T. Selecting Relevant Features for Multi-Label classification of emotions in music. Annual South African Statistical Association Conference, 26-29 November 2018, Johannesburg.
- Van der Merwe, C.J. KNN-Triplot classification with polybags. Annual South African Statistical Association Conference, 26-29 November 2018, Johannesburg.
- Lubbe, S. Comparison of Unfolding and Correspondence Analysis. Annual South African Statistical Association Conference, 26-29 November 2018, Johannesburg.
- Muller, A. Ensemble methods in multi-label classification. Annual South African Statistical Association Conference, 26-29 November 2018, Johannesburg.

SOUTH AFRICAN ACTUARIAL SOCIETY CONVENTION, 24-25 OCTOBER 2018, CAPE TOWN:

- Melville, W., Muller, D. and Van Zyl, N. A projection of the South African government's social security obligations: 2017 – 2037. South African Actuarial Society Convention, 24-25 October 2018, Cape Town.

Presentations at International conferences

CORRESPONDENCE ANALYSIS AND RELATED METHODS (CARME) CONFERENCE, 4-6 FEBRUARY 2019, STELLENBOSCH:

- Rowan, A. and Lubbe, S. Unravelling black box machine learning technique predictions using biplots. Correspondence Analysis and Related Methods Conference, 4-6 February 2019, Stellenbosch.
- Ganey, R. and Lubbe, S. Applications of the principal surface biplot. Correspondence Analysis and Related Methods Conference, 4-6 February 2019, Stellenbosch.
- Nienkemper-Swanepoel, J. le Roux, N.J. and Lubbe, S. To impute or not to impute categorical data: application of multiple correspondence analysis. Correspondence Analysis and Related Methods Conference, 4 – 6 February 2019, Stellenbosch.
- Van der Merwe, C.J. K-nearest neighbour triplot classification with polybags. Correspondence Analysis and Related Methods Conference, 4-6 February 2019, Stellenbosch.

- Le Roux, N.J. and Ngoy, M. What can be learned from representing student performance at the Copperbelt University in Zambia in a biplot? Correspondence Analysis and Related Methods Conference, 4-6 February 2019, Stellenbosch.

CONFERENCE BRICS ON MATHEMATICS, 23-27 JULY 2018, FOZ DO IGUAÇU, BRAZIL:

- Lubbe, S. Visualisation as a tool to understand the sights and sounds of data. Conference BRICS on Mathematics, 23-27 July 2018, Foz do Iguaçu, Brazil.

EUROPEAN CONFERENCE ON DATA ANALYSIS, 4-6 JULY 2018, PADERBORN, GERMANY:

- Nienkemper-Swanepoel, J., le Roux, N.J. and Lubbe, S. Visualising incomplete data with subset Multiple Correspondence Analysis. 5th European Conference on Data Analysis, 4-6 July 2018, Paderborn, Germany.

- Ganey, R. and Lubbe, S. A Biplot based on a Principal Surface. 5th European Conference on Data Analysis, 4-6 July 2018, Paderborn, Germany.

- Schoonees, P. and le Roux, N.J. Computing Neural Reliability from EEG Recordings. 5th European Conference on Data Analysis, 4-6 July 2018, Paderborn, Germany.

- Uys, D.W. Forward stagewise linear regression for ensemble methods. 5th European Conference on Data Analysis, 4-6 July 2018, Paderborn, Germany.

DATA SCIENCE, STATISTICS AND VISUALISATION CONFERENCE, 9-11 JULY 2018, VIENNA, AUSTRIA:

- Ganey, R. and Lubbe, S. A Visualizing Multivariate data using Principal Surfaces Biplots. Data Science, Statistics and Visualisation conference, 9-11 July 2018, Vienna, Austria.

- Le Roux, N.J., Gower, J.C. and Lubbe, S. Can within-group information be used to supplement visual representations of two or three group Canonical Variate Analyses. Data Science, Statistics and Visualisation conference, 9-11 July 2018, Vienna, Austria.

- Uys, D.W. Forward stagewise linear regression for ensemble methods. Data Science, Statistics and Visualisation conference, 9-11 July 2018, Vienna, Austria.

COMPSTAT2018, 28-31 AUGUST 2018, IAŞI, ROMANIA:

- Lubbe, S. A naïve visualisation of Data Science. COMPSTAT2018, 28-31 August 2018, Iaşi, Romania.

- Schoonees, P. and le Roux, N.J. Computing and validating neural reliability from EEG recordings. COMPSTAT2018, 28-31 August 2018, Iaşi, Romania.

- Le Roux, N.J. and Ngoy, M. Biplot perspectives on student performance at the Copperbelt University in Zambia. COMPSTAT2018, 28-31 August 2018, Iaşi, Romania.

INTERNATIONAL CONFERENCE ON MATHEMATICS AND STATISTICS: TEACHING, THEORY AND APPLICATIONS (ATINER), 2-5 JULY 2018, ATHENS, GREECE:

- Viljoen, H. Forecasting Time Series using Stepwise Common Singular Spectrum Analysis or Horizontal Multi-channel Singular spectrum Analysis? 12th Annual International Conference on Mathematics and Statistics: Teaching, Theory and Applications (ATINER), 2-5 July 2018, Athens, Greece.

We hope you enjoyed this edition of the Department of Statistics and Actuarial Science's biannual newsletter. To access all our past newsletters and departmental information visit our website: http://www.sun.ac.za/statistics.

Follow us on social media: https://www.facebook.com/STATISTICSatSUN/

The next newsletter will be distributed in December 2019.