Departement Statistiek en Aktuariële Wetenskap
Department of Statistics and Actuarial Science

Nuusbrief / Newsletter
May / June 2017

In this issue

Department of Statistics and Actuarial Science Celebrates 70th Birthday

Front (left to right): Dr Wiesner Vos, Prof Michael Greenacre, Prof Niël le Roux, Mr Dirk Jooste and Prof Stan du Plessis. Back (left to right): Prof Trevor Hastie, Prof Paul Mostert, Prof Willie Conradie and Mr Wynand Louw.

The Department of Statistics and Actuarial Science celebrated its 70th anniversary on Saturday, 26 November 2016 with a seminar and special lunch.

The seminar covered technical and popular topics presented by young, well-known alumni of the Department as well as international academics with a South African connection. About 100 alumni, staff, colleagues from universities in South Africa and abroad, donors and friends of the Department attended the festivities.

The Department was privileged to host all of its active emeritus professors (Professors Anton Schoeman, Bouwer Bouwer, Caspar Greeff, Mark Classen, Niël le Roux and Tertius de Wet) as well as Professors Nico Laubscher and Daan Nel, past directors of the Statistical Consultation Centre. We are grateful that Stephan and Carl Pretorius, the sons of Prof Faantjie Pretorius who was the first professor in Statistics at Stellenbosch University and the founder of the Department in 1946, were able to join us.

Prof Stan du Plessis (former Dean of the Faculty of Economic and Management Sciences and alumnus of the Department) opened the occasion by welcoming the guests. Prof Willie Conradie then gave a brief overview of the history of the Department. This introduction was followed by three presentations by alumni of the Department. Dr Wiesner Vos, Quant Manager and Statistician in Advanced Measurement Technologies at Google in London, gave an "Overview of statistical analysis and computing at Google"; Mr Dirk Jooste, Proprietary Credit Trader at Investec Bank, discussed "The untimely value of money"; and Mr Wynand Louw, Chief Actuary of Old Mutual, spoke about "How to count in the future…and make a difference".

In the second part of the programme, Emeritus Professor Niël le Roux gave a talk entitled "Living in a 3D environment with multidimensional challenges: Visualising multivariate data".

The programme was concluded with a panel discussion led by Prof Tertius de Wet entitled "Statistics: The road ahead?" Prof Michael Greenacre of Universitat Pompeu Fabra, Barcelona, and Prof Trevor Hastie of Stanford University in the USA gave fascinating insights into the future of our field.

Talks at the 70th Anniversary Celebrations

The 70th anniversary celebrations of the Department of Statistics and Actuarial Science took the form of a seminar with a number of technical though popular lectures presented by young well-known alumni of the Department as well as international academics with a South African connection.

Dr Wiesner Vos

The first speaker was Dr Wiesner Vos. Wiesner completed his Masters degree in Mathematical Statistics at Stellenbosch in 2000, received a DPhil in Applied Statistics from Oxford and has been working at Google in the UK since 2011 as part of an Advanced Measurement Technologies (AMT) group, a research group within Google Ads Engineer. It is the largest statistics team at Google and has 50+ PhD level statisticians spread across four countries. The topic of his talk was "An overview of statistical analysis and computing at Google". In this talk, he discussed his life at Google as a member of the AMT group. Examples of this group's work include:

- The accurate measuring of audiences online;
- Extra reach: How many more people can an advertiser reach if YouTube is added to the TV advertising plan and how effective would this strategy be;
- Attribution: An example of statistical computing at Google of how to determine how much credit each touchpoint should get in an advertising stream for a user.

Wiesner also mentioned that a Google-specific build of R is widely used in the company. This integrates with Google-specific technologies such as BigTable, a Google parallel computing infrastructure. He noted that, in more and more statistical work, Google is switching to Python (invented by a former Googler Guido van Rossum). Python has come a long way: it has become a proper engineering language at Google. Wiesner concluded that TensorFlow provides new opportunities for statistical computation.

Mr Dirk Jooste

The second speaker was Mr Dirk Jooste. Dirk graduated with an MCom in Financial Risk

Management from Stellenbosch in 2006. He is a chartered financial analyst (CFA) holder and currently works as a credit trader at Investec Bank Ltd. The topic of his talk was "The untimely value of money".

Dirk spoke about the past, the present and the future of interest rates and how the current very low interest rate environment is distorting asset valuation techniques. He mentioned that interest rates are at their lowest in 5 000 years. He emphasised that capital markets have become more mobile and, as an emerging market, South Africa is exposed to yield seeking investors (so-called "investment tourists"). He remarked that there is a sense that we've reached the inflection point of these low interest rates.

Mr Wynand Louw

The third speaker of the day was Mr Wynand Louw. Wynand graduated with a BCom Honours degree in Mathematical Statistics from Stellenbosch in 1993 and qualified as a Fellow of the Faculty of Actuaries in 2001. He is currently the chief actuary of Old Mutual Life Insurance Company (South Africa). The topic of his talk was "How to count in the future and make a difference".

The crux of his presentation was that we are entering a new world of information and how we interact with it and what we do with it is crucial. He started with the remark that "age matters, it defines who we are" – different generations have different viewpoints on life.

Wynand emphasised that actuaries are more than just statisticians or finance specialists; they are in the business of promises. As such they are:

- Caretakers, e.g. looking after the promises made to pensioners;
- Narrators of promises, e.g. finance directors should be able to communicate the promises made and the effect of these promises widely throughout the business;
- Dealmakers of promises, e.g. in product design.

He touched on big data and mentioned that the world is changing exponentially. In this changing world, risk is not a new thing; it has always been with us. These include financial risks that endanger the ability to meet financial promises made; human risks such as wars; and environmental risks such as the current water crisis in Cape Town.

Prof Niël le Roux

Emeritus Prof Niël le Roux, a staff member of the Department since 1975 and an internationally recognised expert on visualising multivariate data,

gave a very interesting talk with the title: "Living in a 3D environment with multidimensional challenges: Visualising multivariate data". The theme of his talk was that although we live in a three-dimensional world, it is full of multidimensional challenges. A technique called biplots approximates multidimensional structures and relationships between variables in a two or three-dimensional space. The use of biplots was illustrated with two data sets. First, Niël provided a visualisation of the relationships between the variables from a data set on staff remuneration of Stellenbosch University academics. This data set included a mixture of continuous variables (e.g. age, years' experience and research output score) and categorical variables (e.g. gender and qualification(s)). A second example illustrated how this technique can be used to create visualisations of changes over time. Data on crime in South African provinces between 2004 to 2013 was used and illustrated increases and decreases in certain types of crime in a specific province over the time period.

Prof Tertius de Wet

The morning's programme was concluded with a panel discussion led by Prof Tertius de Wet entitled "Statistics: The road ahead?" In this panel discussion two prominent statisticians discussed their views on Big Data and its influence on the field of Statistics.

Prof Michael Greenacre

Prof Michael Greenacre of Universitat Pompeu Fabra, Barcelona, discussed the democratisation of data – the fact that data is available in abundance to "everyone" over an amazing range and variety of fields. He did warn, however, against the anarchy of data storage with little or no standardisation across the different databases. This makes the pre-processing of the data even more important than in the past with smaller datasets. Furthermore, the data in these databases is often not in its raw form but in some aggregated form, making it more difficult to analyse and draw conclusions.

Prof Trevor Hastie

Prof Trevor Hastie of Stanford University in the USA discussed his own career path as an applied statistician starting at the Medical Research Council, where he worked with small datasets, mostly of a messy form. He still works as an applied statistician, but now with massive datasets, especially in the field of genetics. This meant that he had to become more computationally skilled and also had to start working with, and learning from, computer scientists. He sees an important role for statisticians in data science and says they need to claim it, although they must be willing to learn from other scientists working in data science, especially computer scientists. With the increase in the size of datasets, new techniques to handle issues associated with these needed to be developed. An example of this is the occurrence of wide data that requires a new and different approach to traditional datasets. These present interesting challenges to both applied and theoretical statisticians.

From the input of both speakers, it is clear that statisticians have an important and exciting role to play in Big Data and Data Science, but should also seriously take note of the contributions that scientists from other disciplines have made, and are still making, to this field.

Topical Issues Discussed at the Second Schroders Investment Symposium

A presentation on the very topical subject of integrating political risk into asset allocation kicked off the second investment symposium hosted by global asset management firm Schroders in collaboration with the Department of Statistics and Actuarial Science at Stellenbosch University (SU).

After the success of the inaugural investment symposium hosted last year, representatives from the asset management industry in South Africa gathered in Stellenbosch in March for the 2017 Schroders Institutional Investment Symposium.

This symposium is one of the initiatives flowing from an agreement between SU and Schroders' Multi-asset Investments and Portfolio Solutions (MAPS) division. The agreement also comprises postgraduate bursaries and research collaboration.

International experts discussed the developments and challenges asset allocators are faced with in a rapidly changing industry with industry representatives on Tuesday 28 March and with postgraduate students on Wednesday 29 March.

During her presentation on the impact of political risk Remi Olu-Pitan, fund manager at Schroders, explained that the impact of political shocks is huge when global growth is low.

"We can't hide from political risk – we need to incorporate it into our planning," she said.

Other topics discussed at the symposium included "Using factors to construct portfolios", "Advances in Equity Portfolio Construction", "The risk of mismeasuring risk" and "Putting investment insights into practice".

The Department of Statistics and Actuarial Science's Prof Willie Conradie, who is responsible for coordinating the agreement with Schroders on SU's side, said that it was immensely valuable for academics as well as asset managers from industry to listen to the different speakers and benefit from their knowledge and experience.

"It was also a privilege for our postgraduate students in Financial Risk Management, Actuarial Science and Financial Analysis to learn from and mingle with these international experts," he added.

Prof Wim de Villiers, Rector and Vice-chancellor of SU, opened the symposium and said that the level of interest in this initiative indicates that there is a definite need for collaborations of this nature, bringing academia and the private sector together in the public interest.

He made special mention of the financial assistance in the form of 40 postgraduate bursaries.

“In March, SU concluded the 2016 academic year by awarding a record number of 8339 qualifications. People tend to think that SU is a well-to-do university, but a third of our most recent graduates received some form of financial assistance.”

Prof De Villiers lauded the mutually beneficial partnership model of the collaboration between Schroders and SU, bringing together industry and academic experts who will collaborate on research that both expands academic knowledge and enhances the delivery of investment outcomes.

"It is inspiring to be part of such a forward-thinking initiative," he said.

Gavin Ralston, Head of Official Institutions at Schroders, said that they are very proud of their association with SU. “We very much appreciate the importance of widening access to education. We're very proud of what we've achieved thus far.”

Click here for more information on the presentations.

The financial assistance provided by Schroders in the form of postgraduate bursaries is an important part of this international company's agreement with Stellenbosch University (SU). Here are the 2017 Schroders bursary recipients alongside conference speakers as well as representatives of Schroders, Correlation Risk Partners and SU's Department of Statistics and Actuarial Science.

Attending the conference were (from left to right): Mr Emmanuel Wildiers (Schroders), Dr Ashley Lester (Schroders), Prof Wim de Villiers (Rector, Stellenbosch University), Mr Gavin Ralston (Schroders), Prof Stan du Plessis (Stellenbosch University), Prof Niël Krige (Stellenbosch University), Dr Kevin Kneafsey (Schroders) and Prof Willie Conradie (Stellenbosch University).

Prof Paul Mostert appointed Chairperson of the Department of Statistics and Actuarial Science

Prof Paul Mostert is the new chairperson of the Department of Statistics and Actuarial Science. He succeeds Prof Willie Conradie whose term ended on 31 December 2016.

Prof Mostert studied at the North West University, University of Pretoria and UNISA. He obtained his PhD in Mathematical Statistics from UNISA in 1999. Before joining the Department of Statistics and Actuarial Science at Stellenbosch University in 1997, he was a teacher at Potchefstroom Gymnasium, a subject specialist at Iscor Ltd and a lecturer at UNISA. He was promoted to associate professor in 2002 and also served as a guest professor in survival analysis at the Katholieke Universiteit Leuven in Belgium.

Prof Mostert was a member of the executive committee of the South African Statistical Association from 2005 to 2015 and president of the association in 2014. His research interests are in Bayesian analysis, survival and reliability analysis.

Prof Paul Mostert

Prof Danie Uys

Prof Danie Uys Elected President of South African Statistical Association: 2017-2018

Prof Danie Uys, associated professor in Statistics, was elected President of the South African Statistical Association (SASA) during the annual general meeting of the association in 2016. Prof Uys will be the first president serving a term of two years after an amendment to the constitution in 2015. Prof Uys is the seventh SASA president affiliated to this Department over many years. Previous presidents from this Department were: Profs DEW Schumann (1962), DG Nel (1982), TJ vW Kotze (1984), WJ Conradie (1995), T de Wet (2003) and PJ Mostert (2014).

Award-winning research by Prof Tertius de Wet

Prof Tertius de Wet, Emeritus professor in Statistics, received both the Herbert Sichel medal awarded by the South African Statistical Association (SASA) and the prize for best published research awarded by the Actuarial Society of South Africa for a research paper he co-authored, entitled: "Combining scenario and historical data in the loss distribution approach: A new procedure that incorporates measures of agreement between scenarios and historical data." The Herbert Sichel Medal is awarded annually to the member (or members) of the South African Statistical Association who published the best statistical paper during the previous year. The Actuarial Society of South Africa recognised this paper for the valuable contribution it makes to addressing a practical problem faced by companies in the financial services industry. The article is briefly summarised at the end of this newsletter.

Prof Tertius de Wet (left) received the Herbert Sichel medal from Prof Paul Fatti (chair of the Medal committee) during the 2016 annual SASA conference hosted by UCT.

Department Bids Farewell To Three Staff Members

The Department of Statistics and Actuarial Science said farewell to three staff members in 2016. Prof Nelmarie Louw and Dr Johan van Vuuren retired, while Dr Tom Berning is pursuing a career opportunity in the private sector.

Prof Nelmarie Louw

Prof Nelmarie Louw completed both her undergraduate degree as well as her PhD in Mathematical Statistics at Stellenbosch University. In 2000, she joined the Department as a senior lecturer and in 2003 she was promoted to Associate Professor.

Prof Louw leaves behind a rich legacy in the Department. She was responsible for the important second-year module in Mathematical Statistics and the supervision of a number of Masters students. She also supervised the PhD of one of our current lecturers, Dr M Lamont. Her research fields include classification techniques, kernel methods and multi-label classification. Although these fields are generally acknowledged as particularly challenging areas of research, Prof Louw regularly published in these fields.

Dr Johan van Vuuren

Dr Johan van Vuuren joined the Departement in 1983. In 1998 he completed his PhD, in which he developed techniques to identify outliers and influential observations in multivariate linear regression models. Dr Van Vuuren will be remembered for his creative and innovative approaches to research problems in this field. In the 33 years that Dr Van Vuuren was a member of this Department, he developed a number of service courses, in particular for the Faculty of Engineering, as well as important postgraduate modules in the Applied Statistics stream.

Dr Tom Berning

Dr Tom Berning joined the Department as a lecturer in 2003 and obtained his PhD from Stellenbosch University in 2010. He played an instrumental role in renewing the undergraduate offering in Applied Statistics and was also very involved with the South African Statistical Association, where he fulfilled the role of chairperson of the Western Cape for many years.

Not only will the shoes of these three colleagues be a challenge to fill, but their daily friendship will also sorely be missed.

Dr Tom Berning

From left to right: Prof Willie Conradie, Dr Johan van Vuuren, Prof Nelmarie Louw and Prof Stan du Plessis

Department of Statistics and Actuarial Science Appoints Two New Lecturers

Dr David Hofmeyr, who completed his PhD at Lancaster University and graduated in 2016, joins the Department as a senior lecturer in Statistics.

- **What would you like to offer students?** I believe that for many people, myself included, their understanding of statistics and probability can be traced back to a few moments of comprehension; be they from an expert explanation from a teacher, or from personal epiphany. What I really hope to be able to offer students is a learning environment that promotes these moments; either by engaging students in the material to the extent that they have their own moments of realisation or by sharing with them my own experiences and understanding.
- **Please name some of your career highlights:** I was awarded the First Rand/Laurie Dippenaar scholarship for international postgraduate study in 2010, which took me to the University of Edinburgh to study for an MSc in Operations Research. After that I received a studentship (PhD scholarship) from Lancaster University to study for a second masters, and subsequently a PhD, within the STOR-i Centre for Doctoral Training. I received the Kingsman Prize in 2016, which is awarded to the top doctoral researcher in the Department of Management Science at Lancaster University. In the final year before joining Stellenbosch I held a postdoctoral research fellowship within the STOR-i Centre for Doctoral Training.
- **What skills taught in Mathematical Statistics do you think is particularly valuable in the "real world"?** Technological advancements have made the collection and storage of large amounts of data almost trivial. As a result, in almost all fields, decision-making is increasingly based on the results of robust analyses of these data. Statistics forms the basis for extracting useful and relevant information from data, and so it is no surprise that statistics is being heralded as one of the most important and saleable skills currently.
- **How do you plan to assist students to acquire these skills?** A skilled statistician must be well versed both in the mathematical modelling of real world situations within a statistical framework, and, perhaps more

importantly, the interpretation of the results obtained. I hope to be able to offer students both of these skills and stress the importance of utilising them in tandem.

- **Why would you encourage prospective students to come to Stellenbosch University?** The Department of Statistics and Actuarial Science at Stellenbosch University offers an extremely diverse set of skills to prospective students. The staff and researchers have exemplary skills in classical and modern statistics; both in the applied and methodological sides of the discipline.
- **What do you think are the important characteristics that a (prospective) Mathematical Statistics student should possess if s/he wants to pursue a career in Mathematical Statistics?** A sound mathematical understanding is essential for the study and implementation of statistical methods, and is a prerequisite for any prospective student. On top of that, an ability to translate real world situations into a mathematical framework, and vice versa, is vital.
- **What do you enjoy most about Mathematical Statistics?** My passion is in mathematics, and what I love about statistics is being able to not only engage

with this passion, but also have it be made relevant to real world problems by the importance statistics has in decision-making.

- **What are the opportunities in Mathematical Statistics?** Data is everywhere, quite literally. The opportunities available to those who are data literate are too numerous to list. Almost every company and organisation nowadays makes use of the analysis of data in their decision and policy-making. People have realised that if the vast amounts of information which can be extracted from data are ignored, it is extremely difficult to remain competitive, or even relevant.
- **In your opinion, what should the Department of Statistics and Actuarial Science look at to make sure they're keeping up with industry, if anything?** I think the most important thing is to establish and maintain a dialogue between academics and industry, so that they remain on the same page. Many industrial organisations are lagging in the methods they are using as they are not staying up to date with the state of the art. At the same time, however, unless academics are aware of the problems organisations are facing in the now, there is a risk that the research focus is not appropriately directed at the most relevant issues.

***

Prof Sugnet Lubbe completed both her undergraduate and postgraduate degrees at Stellenbosch University, obtaining her PhD in 2001. She joins the Department as Professor in Statistics.

- **What would you like to offer students?** Insight into the exciting world of applying mathematics to tell the story of the data.
- **Please name some of your career highlights:** I spent 13 years in industry as researcher statistician / senior statistician before moving to academia where I was associate professor at UCT from July 2009 to December 2016.
- **What skills taught in Statistics do you think is particularly valuable in the "real world"?** How to think analytically about problem-solving.
- **How do you plan to assist students to acquire these skills?** At different levels students are presented with different problems to solve – in second year, mastering the concepts of estimation and hypothesis testing, at masters and PhD level, tackling a multivariate research problem.

- **Why would you encourage prospective students to come to Stellenbosch University?** It is not only an academically outstanding university but also a university with a unique culture.
- **What do you enjoy most about Statistics?** "Getting to play in everyone else's backyard" (John Tukey). I can use my skills and knowledge to apply statistics to many other interesting fields ranging from health science, sport and finance to antique furniture.
- **What are the opportunities in Statistics?** As one of the most sought after jobs internationally, the opportunities for data scientists are growing by the day.
- **In your opinion, what should the Department of Statistics and Actuarial Science look at to make sure they're keeping up with industry, if anything?** We are in the process of looking into a specialised data science offering at undergraduate level and also to expand our postgraduate offerings in big data and data science.
- **Why did you choose to work at Stellenbosch University?** It is a university with a very good reputation locally and internationally and we are not stagnating but keeping up with new developments and changes over time.

Graduation 2016/17: PhD And Masters Students

Name of student: Retha Luus
Degree: PhD (Statistics)
Supervisor: Prof T de Wet
Title of thesis: Statistical inference of the multiple regression analysis of complex survey data

Abstract:
Complex survey sampling methodology formed the main research area with sub-areas under this including cross-validation, bootstrap, multilevel modelling and sampling weight trimming. The dissertation's main objectives were investigating the effect of ignoring the complex survey design when modelling such data; the importance of correctly calculating final sampling weights; the effect of trimming the weights on inference precision; and using resampling methods such as the bootstrap and jackknife for variance estimation of the estimators of model coefficients. The contributions from this dissertation included the development of a model to simulate such data for the evaluation of new methodologies under controlled conditions; the extension of model evaluation methods for application to survey-weighted linear models; non-parametric linear modelling as an alternative to parametric modelling as well as using the Hill estimator as extreme sampling weight cut-off identification method. Extensive simulation studies, using both simulated as well as real-world data, confirmed that modelling survey data incorrectly leads to distorted results; model evaluation methods show promise when applied to survey-weighted modelling; non-parametric modelling often outperformed the alternatives and Hill weight trimming mostly improved the precision of model parameter estimators in comparison to the untrimmed alternatives.
Find link to thesis here.

Name of student: Benjamin Gurr
Degree: MCom (Statistics)
Supervisor: Prof NJ le Roux
Title of thesis: An application of geometric data analysis techniques to South African crime data

Abstract:
South Africa is recognised as having the eighth highest homicide rate in the world; along with a generally high level of violent crime, the analysis is conducted on reported violent crime statistics in South Africa. This study explores the application of several types of geometric data analysis methods to the study of crime in South Africa. These include: correspondence analysis, the correspondence analysis biplot, and the log-ratio biplot. The emphasis is on visualising the violent crime statistics in South Africa for the 2004-2013 reporting period, using the definition of the various crimes as reported by the South African Police Services.
Find link to thesis here:

Name of student: Anton Scharnick
Degree: MCom (Financial Risk Management), Cum Laude
Supervisor: Prof W Conradie
Title of thesis: Dynamic hedging of with-profit annuities in the presence of transaction costs

Abstract:
Insurance companies, asset managers and pension funds that manage with-profit annuity policies are exposed to market risk. This risk can be managed with the use of dynamic hedging. Dynamic hedging can, however, only completely remove the market risk if hedging occurs on a continuous basis, and if no transaction costs are involved. In practice, the hedger is therefore faced with the challenge to find the appropriate time to hedge. The purpose of this study is to compare the performance of the different hedging strategies that are used in practice, principally in the case of hedging with-profit annuity policies and to investigate whether a more optimal hedging strategy exists than what is currently being used by most insurers.

Name of student: Anri Kotze
Degree: MCom (Financial Risk Management)
Supervisor: Prof W Conradie
Title of thesis: A detailed analysis of credit risk models

Abstract:
The financial crisis of 2008 and the events following it have highlighted the importance of measuring and managing credit risk properly. In order to manage credit risk effectively, it is essential that the importance of the Probability of Default (PD) input parameter and its estimation are understood. Several models for estimating PDs have been introduced. These models differ in distributional assumptions, restrictions, calibration and application. With this background in mind, the purpose of this study is to review the literature on PD models as well as industry implementation both in the corporate and retail market. The main focus is on the Structural Models introduced by Merton.

Name of student: Margaret de Villiers
Degree: MCom (Mathematical Statistics)
Supervisor: Prof D Uys
Title of thesis: Predicting tomato crop yield from weather data using statistical learning techniques

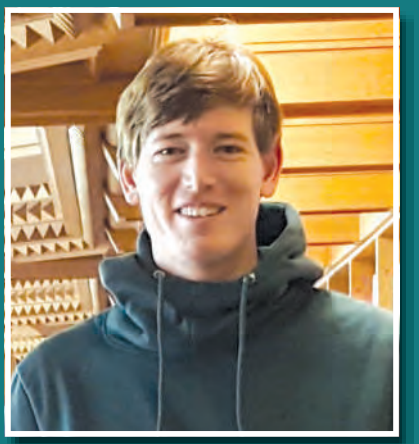
Abstract:
This project involved developing statistical models for predicting harvest quantity of field-grown tomato crops using crop and weather data. The weather variables consisted of summary statistics of the readings over the growing period of each crop. Median and total harvest density (t/ha) were each modelled using statistical learning theory techniques. The median and total harvest density models developed had estimated mean absolute prediction errors of 0.37 t/ha and 12.67 t/ha, respectively. The large prediction errors are due to the absence of important predictors and the summary of the weather time series over the crops' growing periods into single values.
Find link to thesis here.

Name of student: Arnu Pretorius
Degree: MCom (Mathematical Statistics), Cum Laude
Supervisor: Dr S Bierman
Co-supervisor: Prof SJ Steel
Title of thesis: Advances in random forests with application to classification

Abstract:
Leo Breiman's popular random forest algorithm (Forest-RI), and related ensemble classification algorithms which followed, formed the focus of this study. Some of the contributions of the study included the following: a review of random forest algorithms that were developed since the introduction of Forest-RI, including a novel taxonomy of random forest classification algorithms; an investigation of possible decomposition of the expected prediction error of ensemble classifiers into bias and variance components; an empirical study of ensemble learners, including bagging, boosting and Forest-RI; a novel random forest framework, viz. oblique random rotation forests, which serves as an example of a heuristic approach towards novel proposals based on bias-variance analyses; and a meta-analysis of all research papers on Forest-FI and related ensemble classifiers, leading to a novel two-step procedure allowing comparison of multiple algorithms over multiple data sets and across several papers.
Find link to thesis here.

Name of student: Monika du Toit
Degree: MCom (Mathematical Statistics)
Supervisor: Prof SJ Steel
Title of thesis: L-Classifier chains classification and variable selection for multi-label datasets

Abstract:
Multi-label classification extends binary and multi-class classification to scenarios where every data case is assigned several labels simultaneously. Applications include labelling images with tags and identifying instruments that are playing together in a musical piece. Variable selection is important in multi-label data analysis. In this thesis, a multi-label classification approach called L-classifier chains (LCC) is proposed. This method implements a compromise between simple classifier chains and the ensemble of classifier chains procedures. The results from LCC include multi-label predictions and a matrix of variable importance values. Application of the LCC classifier is illustrated using benchmark multi-label datasets, simulated datasets and a practical dataset obtained from a South African credit bureau. Find link to thesis here.



Name of student: Sven Buitendag
Degree: MCom (Mathematical Statistics), Cum Laude
Supervisor: Prof T de Wet
Title of thesis: The saddle-point method and its application to the Hill estimator

Abstract:
The theory underlying the saddle-point approximation is discussed and illustrated with an application to approximate the distribution of the Hill estimator in Extreme Value Theory. The saddle-point approximation can be applied to approximate the distribution of any random variable with a well-behaved cumulant generating function, with highly accurate results.
Find link to thesis here.

Name of student: Chané Orsmond
Degree: MCom (Mathematical Statistics), Cum Laude
Supervisor: Prof SJ Steel
Title of thesis: Statistical classification procedures for analysing functional data

Abstract:
Functional data are obtained through measuring variables at a set of discrete evaluation points over a continuum such as time or wavelength. Functional extensions of traditional statistical methods are considered in this research. Linear discriminant analysis for functional data and functional support vector machines are investigated. To address the high correlations amongst the inputs of a functional data set, the fused lasso is discussed. A sparse equivalent of partial least squares is also investigated. An infrared spectroscopy data set is considered for practical implementation of the fore mentioned functional data analysis techniques. The procedures are compared in terms of classification accuracy and variable selection properties.
Find link to thesis here.

Graduation Joy

Several Schroders bursary recipients graduated in 2016. Here are KJ Sebitlo, Prof Willie Conradie, Simonn Benjamin and Justin Perrang.

2016 Final-year prize winners

At the end of each year, the Department awards prizes to the best final-year undergraduate students in each of the four disciplines offered by the department. At a cocktail function for all the final-year students, Prof Ronel du Preez, Vice-Dean: Teaching of the Faculty of Economic and Management Sciences, handed over cash prizes to the following students:

Actuarial Science: Gerben Draaijer

Gerben matriculated from Paul Roos Gymnasium. He is currently enrolled for an Honours degree in Actuarial Science.

Why did you decide to study Actuarial Science? I had an interest in maths and economics so it seemed the logical choice. I also heard rumours that it's a much nicer job than being an accountant.

Do you think Actuarial Science is a useful subject that can bring change to our society at large? I think the actuarial profession is already providing a useful service to society, but yes, I do believe there is probably room for more innovation with regards to social matters.

What are your future career plans? I hope to work somewhere in Cape Town (it's nice and close to the sea and the mountains) and at the moment consulting seems like one of the most attractive options.

What are some of your hobbies / other interests? Anything related to being outdoors like running and mountain biking. More generally any sport that can be played with friends like cricket, tennis or touchies. I also enjoy playing guitar in my free time.

Statistics: Connor McCann

Connor matriculated from Brackenfell High School. He is currently enrolled for an Honours degree in Financial Analysis in the Department of Business Management.

Why did you decide to study Statistics? Statistics offers the ability to examine the real world and determine whether something is random or has an actual relationship. I most liked the application of theoretical models to the real world, using statistical programs.

Do you think Statistics is a useful subject that can bring change to our society at large? Statistics can be applied to practically any situation or field and can thus be used to positively influence society at large through analysing situations to determine helpful relationships that can be utilised to optimise and better society.

What are your future career plans? My plan is to go into the investment world, specifically to become an asset / portfolio manager and analyst.

What are some of your hobbies / other interests? I enjoy watching sport, specifically soccer and tennis, playing table tennis, and watching TV series and movies. I also try to keep up to date with world events and politics.

Mathematical Statistics: Ian Louw

Ian matriculated from Paul Roos Gymnasium. He is currently enrolled for an Honours degree in Mathematical Statistics.

Why did you decide to study Mathematical Statistics? I tend to enjoy anything mathematical and challenging, but I also like to apply my knowledge to real world problems. In Mathematical Statistics I found a degree that combines all three of these areas. What I like most is how applicable the subject is in everything we do.

What do you like most about the Department and your postgraduate degree? The Department: The staff are exceptional and they stand out from the crowd. I don't think they are appreciated enough. The postgraduate degree: The practical approach to learning.

Do you think Mathematical Statistics is a useful subject that can bring change to our society at large? Surely. The vast (and ever increasing) amount of data being generated and collected requires experts in this field to make sense of it. Studying Mathematical Statistics will give you the statistical tools to perform such analyses but also give you the necessary knowledge to develop new

models, frameworks and theories to overcome modern types of problems.

What are your future career plans? I have a passion for the risk management space and I would like to use this passion in the investment world.

What are some of your hobbies / other interests? I'm an avid golfer and very interested in any kind of sport. Coffee and latte art is another hobby of mine and I would like to obtain my Barista qualification.

Financial Risk Management: Melissa Kemp

Melissa matriculated from Stellenberg High School. She is currently enrolled for an Honours degree in Financial Risk Management.

Why did you decide to study in this Department? My undergraduate years were not enough. I wanted to gain more knowledge in this field. I've always been attracted to the corporate world – the people, the problems, the management of wealth – and I think pursuing my Honours and possibly Masters will help me rise above my competition and become a force to be reckoned with.

What do you like most about the Department and your postgraduate degree? Apart from the personal level of teaching we receive, our programme provides us with a broad exposure to what to expect in the real world. They go above and beyond to help us build the stepping stones of our future. The greatness of the Department is due to the world-class lecturers.

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

Yes. This subject helps you develop various skills to better understand problems and find innovative solutions. I believe that knowledge about risk and the management thereof can assist with a vast number of the economic issues that we are currently faced with.

What are some of your hobbies / other interests? When I am not studying you can find me at a music festival, the beach or kicking it back at a braai.

Student Prize Winners: 2016

Students Benefit from Essay Competition

As part of a larger collaboration between Stellenbosch University and Schroders, 10 students in the Faculty of Economic and Management Sciences walked away with R20 000 each after participating in an essay competition.

This essay competition forms part of a joint investment research initiative that stems from an agreement signed between the Department of Statistics and Actuarial Science and Schroders.

The competition was open to postgraduate students in the Departments of Statistics and Actuarial Science, Business Management and Economics on topics proposed by Schroders and its strategic partner Lombard Insurance. Each topic was assigned to a mentor by Schroders, who were available via Skype to give advice and practical insight.

Ashley Lester, Global Head of Multi Asset and Portfolio Solutions Research (MAPS) at Schroders, reviewed all the essays. "I am particularly happy to note the high quality of many of the entries. Selecting the 10 winners was not an easy task, and I fully anticipate that it will become even more difficult in future years, as students become better acquainted with the essay competition," he said.

Carel van der Merwe of the Department of Statistics and Actuarial Science, who managed and organised the competition with Schroders, believes the essay competition is a great opportunity for students to research subjects relevant to industry and to connect and network with senior staff members at respected and well-known companies.

The winners were Coetzee Janse van Rensburg, Peter-John Clift, Peru Govindasamy, Rina de Villiers, Jannes Germishuys, Anri Kotzé, Gcinile Dlamini, Nadia Burger, Rebecca Selkirk and Johann Pfitzinger.

Pfitzinger, who achieved the highest mark, enjoyed working on the topic he chose ("The effects of unconventional monetary policies").

"It is particularly rewarding when the work receives such generous recognition," he added.

Prize winners and roleplayers: Back (left to right) Johann Pfitzinger, Peru Govindasamy, Peter-John Clift, Prof Stan du Plessis (former Dean of the Faculty), Coetzee Janse van Rensburg, Mr Doug Abbott (Schroders), Jannes Germishuys, Prof Willie Conradie (SU), Prof Neil Krige (Chairperson of SU's Development Office) and Carel van der Merwe (SU). Front (left to right) are Rina de Villiers, Gcinile Dlamini, Nadia Burger, Rebecca Selkirk and Anri Kotzé. Photo credit: Hennie Rudman

SAS® Honours Competition 2016

Each year the South African Statistical Association (SASA) awards monetary prizes to students who participate in the annual SAS sponsored honours project competition. The winner of the competition is also invited to attend the annual SASA conference to present the winning project. For a second year in a row, this prestigious prize was awarded to an Honours student from our Department. In 2015 the prize went to Margaret de Villiers and Dalene Bezuidenhoud. The 2016 prize was awarded to Luca Steyn (currently registered for a Masters in Mathematical Statistics). The title of the honours project is: "Statistical Modelling with Quantile-based Methods and L-moments." He received the prize from Mr Murray de Villiers, manager of the global academic programme at the SAS Institute during the 2016 annual SASA conference.

Bursaries and Grants: 2016

SASA-NRF Academic Statistics Grants

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.

Administered by the NRF and the South African Statistical Association, the fund is used to provide full-time bursaries to Masters and PhD students, as well as grants to postdoctoral students and postgraduate supervisors. The only criterion to obtain a bursary is academic merit.

Stellenbosch University's first Masters cohort (2015/2016) all graduated last year. They are: Margaret de Villiers, Chané Orsmond and Sven Buitendag.

The second Masters cohort (2016/2017) are busy with the research component of their degree. They are Luca Steyn, Tessa Rodseth, Jandré Marais and Frances Coetzer.

The following students received bursaries for 2017:
Masters: Annegret Muller and Dylon Botha.
PhD: Sven Buitendag.

Back (from left to right) are Dylon Botha, Luca Steyn, Tessa Rodseth and Jandré Marais, and front (from left to right) are Annegret Muller and Frances Coetzer.

Faantjie and Lettie Pretorius Bursary for Undergraduate Students

Prof Faantjie Pretorius was the first professor in Statistics at Stellenbosch University. His heirs donated a substantial sum in his memory and the return on this investment is used to provide bursaries to deserving undergraduate students of the Department every year. In 2017, Suveer Nundlall and Sibulelo Mfanta shared the bursary from this fund.

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 while three members of their senior management team have obtained postgraduate qualifications in Financial Risk Management, also from the Department of Statistics and Actuarial Science. 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 eight third-year students in Financial Risk Management.

On the photo are six of the eight bursary recipients. From left to right are Ayesha Abrahams, Jean Naude, Delia Sandilands, David Rodwell, Amber Newton and Robert Cronje. Ashleigh Miller and MMadiawetsa Mahwiliri were absent.

Stellenbosch Alumni Become Fellow Members Of Actuarial Society

The road to qualifying as an actuary is a long one. Not only must students pass a number of demanding examinations set by the Actuarial Society of South Africa, but also accumulate relevant practical work experience before being admitted as Fellows of the Actuarial Society of South Africa (FASSA). Qualifying as a Fellow is therefore a great achievement. The following alumni from Stellenbosch University were admitted as Fellows in 2016 and the Department would like to congratulate them: RA Brown, JF de Villiers, S du Plessis, JvO Fourie, RF Jacobs, JJM Kriel, EJ le Roux, JJ Liebenberg, WA Louw, R Meyer, J Nicholas, H Strauss, J van der Merwe and PJA van der Merwe.

Combining Scenario And Historical Data In The Loss Distribution Approach: A New Procedure That Incorporates Measures Of Agreement Between Scenarios And Historical Data

PJ de Jongh, T de Wet, H Raubenheimer and JH Venter

Many banks worldwide currently use various versions of the so-called loss distribution approach (LDA) to calculate Value-at-Risk (VaR) for Operational Risk under the Basel Accord's Advanced Measurement Approach (AMA).

One of the important building blocks of the LDA approach is the annual aggregate loss distribution in a particular unit of measure or operational risk category (ORC). According to the Basel II Accord, the 99.9% VaR of the aggregate loss distribution is used as a measure of the regulatory capital (EC) in that particular ORC.

Many researchers have highlighted several deficiencies and expressed concern about elements of the LDA approach, in particular the estimation of the severity distribution and the subsequent estimation of an extreme VaR of the aggregate loss distribution. Common sense suggests that if historical data are to be used, at least a 1000 years of loss data for the ORC in question would be required to estimate such a quantity with reasonable accuracy. Typically banks have between 5 and 10 years of internal loss data available. The Basel Accord suggests the use of scenario assessments to improve severity distribution estimation.

Various approaches to incorporating scenario assessments into the modelling of the severity distribution have been proposed in the literature. A popular method is to model the body and tail of the severity distribution separately – using the empirical distribution to model the body and the Generalized Pareto Distribution (GPD) to model the tail.

In our paper the focus was to investigate another approach, referred to as the 1-in-c years scenario approach. In this approach the scenario makers are asked to answer the question: 'What loss level, q_c ,is expected to be exceeded once every c years?' The AMA models of banks suggested that popular choices for c vary between 5 and 100 and often 3 values for c are used. As an example, one bank used $c = 7$, 20 and 100 and motivated the first choice as the number of years of historical data available to them. In this case the largest loss in the historical data may serve as a guide for choosing q_7 since this loss level has been reached once in 7 years. The other choices of c are selected in order to obtain a scenario spread within the range that one can expect reasonable improvement in accuracy from the experts' inputs.

The proposed method consisted of estimating the severity distribution from a combination of an estimated underlying distribution, e.g. a Burr with estimated parameters, and measures of agreement between the historical data and the scenario assessments.

Specifically, let $F(x;\theta)$ denote a suitable family of distribution functions to model the severity distribution. Fit $F(x;\theta)$ to the historical data and denote the parameter estimate by $\hat{\theta}$. We assumed that scenario assessments $\tilde{q}_7$, $\tilde{q}_{20}$ and $\tilde{q}_{100}$ are available and denoted by $\hat{p}_7 = 1 - 1/(7\hat{\lambda})$, $\hat{p}_{20} = 1 - 1/(20\hat{\lambda})$, $\hat{p}_{100} = 1 - 1/(100\hat{\lambda})$ the estimates of the corresponding true probabilities $p_7 = 1 - 1/(7\lambda)$, $p_{20} = 1 - 1/(20\lambda)$, $p_{100} = 1 - 1/(100\lambda)$. Here λ and $\hat{\lambda}$ denote respectively the parameter of the frequency distribution and its estimator. The agreement ratios are defined as

$$R(7) = \frac{\hat{p}_7}{F(\tilde{q}_7;\hat{\theta})}, \quad R(7,20) = \frac{\hat{p}_{20}-\hat{p}_7}{F(\tilde{q}_{20};\hat{\theta})-F(\tilde{q}_7;\hat{\theta})},$$

$$R(20,100) = \frac{\hat{p}_{100}-\hat{p}_{20}}{F(\tilde{q}_{100};\hat{\theta})-F(\tilde{q}_{20};\hat{\theta})} \text{ and } R(100) = \frac{1-\hat{p}_{100}}{1-F(\tilde{q}_{100};\hat{\theta})}.$$

Clearly if our estimates were actually exactly equal to what they are estimating, these ratios would all be equal to 1.

The method then estimates the severity distribution function G by an adjusted form of $F(x,\theta)$ denoted by $\tilde{H}$ and defined as follows:

$$\tilde{H}(x) = \begin{cases} R(7)F(x;\hat{\theta}) & for\ x \le \tilde{q}_7 \\ \hat{p}_7 + R(7,20)[F(x;\hat{\theta}) - F(\tilde{q}_7;\hat{\theta})] & for\ \tilde{q}_7 < x \le \tilde{q}_{20} \\ \hat{p}_{20} + R(20,100)[F(x;\hat{\theta}) - F(\tilde{q}_{20};\hat{\theta})] & for\ \tilde{q}_{20} < x \le \tilde{q}_{100} \\ \hat{p}_{100} + R(100)[F(x;\hat{\theta}) - F(\tilde{q}_{100};\hat{\theta})] & for\ \tilde{q}_{100} < x < \infty \end{cases}$$

In assessment of the method based on real and simulated data, the method performed very well.

Academic Articles – Published Or Presented In 2016

Published Articles

- BEIRLANT J, BARDOUTSOS A, DE WET T. Bias reduced tail estimation for censored Pareto type distributions. Statistics and Probability Letters 2016; 109:78-88.
- CAEIRO F, GOMES I, BEIRLANT J, DE WET T. Mean-of-order p reduced-bias extreme value index estimation under a third-order framework. Extremes 2016; 19:561-589.
- DE WET T, GOEGEBEUR Y, GUILLOU A, OSMANN M. Kernel regression with Weibull-type tails. Annals of the Institute of Statistical Mathematics 2016; 68:1135-1162.
- DE WET T, Ã–STERREICHER F. Fisher's information and the class of f-Divergencies based on extended arimoto's entropies. South African Statistical Journal 2016; 50:43-64.
- GOWER J, LE ROUX NJ, LUBBE S. Biplots: Qualititative data. Wiley Interdisciplinary Reviews: Computational Statistics 2016; 8:82-111.
- MOSTERT PJ, BEKKER A, ROUX J. A Bayesian analysis for censored Rayleigh model using a generalised hypergeometric prior. South African Statistical Journal 2016; 50:195-219.
- PEPLER PT, UYS DW, NEL DG. A Comparison of some methods for the selection of a common eigenvector model of the covariance matrices of two groups. Communications in Statistics - Simulation and Computation 2016; 45:2917-2936.
- SCHOONEES PC, LE ROUX NJ, COETZER LJ. Flexible Graphical Assessment of Experimental Designs in R: The vdg Package. Journal of Statistical Software 2016; 74(3):10.18637/jsss.v074.i03.
- STRYDOM ML, CORUBOLO D, NEL C. Changes in mortality of people living with HIV in South Africa and their potential implications for life assurers. South African Actuarial Journal 2016; 16:1-35.
- VAN HEERDEN JD, VAN RENSBURG P. The impact of liquidity on the cross section of equity returns on the Johannesburg securities echange. Economics, management and financial markets 2016; 11:59-86.
- VAN ZYL N, VAN ZYL DJJ. The impact of behavioural economics and finance on retirement provision. South African Actuarial Journal 2016; 16: 91-125.

Presentations

- Van Zyl, N and Van Zyl, DJJ. The impact of behavioural economics and finance on retirement provision. Presented at the Pensions, Risk and Investment Conference 2016 with AFIR/ERM in Edinburgh, United Kingdom. June 2016.
- Mostert, P.J. & Van Rooyen, R. Approximate objective Bayesian analysis for the generalised compound Rayleigh distribution. International Society for Bayesian Analysis Conference (ISBA2016), Cagliari, Sardinia, Italy, June 2016.
- Van der Merwe, C. Financial reporting for quantitative experts. Gent University.
- Van der Merwe, C. Approximating risk-free curves in sparse data environments. Nelson Mandela Metropolitan University.
- Pretorius, A., Bierman, S., and Steel, S.J. A Meta-Analysis of Research in Random Forests for Classification. Proceedings of Pattern Recognition Association of South Africa and Robot and Mechatronics International Conference (PRASA-RobMech, 2016)

58th Annual conference of the South African Statistical Association (SASA 2016), Cape Town, South Africa

- Viljoen, H. Investigating Stepwise Common Singular Spectrum analysis and Multi-channel Singular Spectrum Analysis.
- Bierman, S. and Steel, S.J. A Multivariate Regression Approach to Multi-Label Classification.
- Pretorius, A., Bierman, S. and Steel, S.J. A Bias-Variance Analysis of Ensemble Learning for Classification.
- Harvey, J. and van der Merwe, A.J. The Performance of Non-Informative Priors for the Bivariate Lognormal Distribution.
- Contardo-Berning, I., Beukman, E. and Ward, C. Statistical Process Control and the enhancement of water quality in the Western Cape.
- Le Roux, N.J. and Bakk, Z. Biplot-based visualizations of latent class models.
- Luus, R., Neethling, A. and de Wet, T. Simulating Complex Sampling Data for the evaluation of Survey-Weighted Linear Models.
- Neethling, A., Luus, R. and de Wet, T. The Effectiveness of Weighting and Bootstrap in the Estimation of Model Parameters under Complex Sampling.
- Nienkemper-Swanepoel, J., Le Roux, NJ and Lubbe, S. Multidimensional graphical comparisons of imputation methods for missing values in categorical data sets.

- Nienkemper-Swanepoel, J., Le Roux, N.J., and Lubbe, S. Visualizations of multiple imputations using generalized orthogonal procrustes analysis. COMPSTAT 2016, Oviedo Spain.
- Le Roux, N.J., Gower, J., and Lubbe, S. Consequences of combining Fisher's optimal scores with biadditive models. 22nd International Conference on COMPUTATIONAL STATISTICS (COMPSTAT 2016), Oviedo Spain.
- Le Roux, N.J., Coetzer, R., and Rossouw, R. – Multivariate Statistical Process Evaluation and Monitoring for Complex Chemical Processes. International Conference on Information Complexity and Statistical Modeling in High Dimensions (IC-SMHD-2016), Cappadocia, Nevsehir, Turkey.