DEPARTMENT OF STATISTICAL SCIENCES

HONOURS PROGRAMMES 2022

MODULE INFORMATION

Note: All outlines are provisional and subject to change in response to university policy around teaching and examination during the pandemic.

Contents

Advanced Probability Theory (2020)

Synopsis

The course begins by covering some probability theory topics that were not covered in STA3045F, before moving on to study advanced topics in martingale theory. The course then introduces continuous-time stochastic processes and general stochastic integration. Some applications involving Brownian motion are given and two important theorems (Feynman-Kac and Martingale Representation) are proved. The last part of course is dedicated to studying Levy processes.

Note: this module is a continuation of the Probability Theory section of STA3045F and knowledge of that material will be assumed. Students interested in the module are strongly advised to revise STA3045F and Real Analysis from MAM2000W.

Course Content and Lecturer Information

Lecturer: Mr Melusi Mavuso Room 5.64 Melusi.Mavuso@uct.ac.za

Topics:

1. Product spaces
2. Convergence of sequences of random variables
3. Martingale convergence and optional sampling theorems
4. Continuous-time σ-algebras: optional, previsible, and progressive.
5. The Doob-Meyer decomposition theorem and quadratic variation
6. Square integrable martingales
7. Stochastic integration
8. Local martingales, semimartingales
9. Feynmac-Kac and Martingale Representation
10. Levy processes

Relevant Texts:

1. *Probability* – A. N. Shiryaev
2. *Probability with Martingales* – D. Williams
3. *Brownian Motion and Stochastic Calculus* – I. Karatzas and S. Shreve
4. *Diffusions, Markov Processes and Martingales* vol 1 and 2 – L.C. Rogers and D. Williams
5. *Limit Theorems for Stochastic Processes* – J. Jacod and A. N. Shiryaev
6. *Theory of Martingales* – R. Lipster and A.N. Shiryaev
7. *Statistics of Random Processes* volume I: General Theory – R. Lipster and A.N. Shiryaev
8. *Levy Processes and Stochastic Calculus* - D. Applebaum

Assessment

- **Composition**: Assessment will consist of a single class test and exam.
- **Format**: Both the class test and exam will be open-book.
- **DP Requirement**: A minimum of 35% for the class test plus attendance of lectures.
- **Final Mark**: The final grade for the course will be based on:

Class Test:	30%.
Final Exam:	70%.

Applied Spatial Data Analysis

Synopsis

Applied Spatial Data Analysis is designed to introduce you to statistical techniques designed to analyse spatial data where location of observations is important. The course will cover handling spatial data in R, visualising spatial data, analysing three different types of spatial data, namely geostatistical, point pattern and lattice data. Applications will be drawn from a wide variety of fields in the natural, business and social sciences. Emphasis will also be placed on the interpretation of computer generated results using R.

Lecturer Information

Lecturer:	Dr. Sebnem Er	Room 5.55	sebnem.er@uct.ac.za
Lecturer:	Mr. Sulaiman Salau	Room 5.54	sulaiman.salau@uct.ac.za

Course Content

- Interpolation and Geostatistics
  - Estimating the spatial correlation: the variogram
  - Spatial prediction
  - Model diagnostics
  - Geostatistical simulation
  - Further methods for analysis of geostatistical data
- Point Patterns
  - Exploratory analysis of spatial point pattern data
  - Spatial randomness tests
  - Ripley's K-Function
  - Analysis of spatial random processes
  - Further methods for analysis of point patterns
- Lattice Data
  - Spatial neighbours and weights
  - Spatial autocorrelation
  - Modeling spatial lattice data
  - Further methods for analysis of lattice data

Assessment

- **Format**: Assignments will consist of a formal write-up to be submitted in PDF format on Vula via the assignments tab. Page limits will be announced upon release of the assignment. Where prediction sets are required, these will be handed in electronically via Vula along with the assignment. Marks will be allocated for the quality of the write up.
- **DP Requirement**: Completion of all projects and tasks by the due dates..
- **Final Mark**: The final grade for the course will be based on:

Online quizzes:	10%.
Participation:	5%.
Presentation:	15%.
Assignment 1:	30%.
Assignment 2:	40%.

Analytics

Synopsis

This course will cover computationally intensive statistical methods for analysing datasets of various sizes, with applications covering parallelisation of code and high performance computing. The methods covered include approaches to both regression and classification problems in a supervised learning context, as well as unsupervised learning methods. The course consists of three broad sections: Tree-based Methods, Artificial Neural Networks, and Unsupervised Learning. The first section covers regression and classification trees, bagging, random forests and boosting. The second section will introduce students to the basic principles of artificial neural networks, the concepts of stochastic gradient descent and backpropagation, and various network architectures and their uses. The last section will cover statistical methods for classifying observations into groups where the group memberships of the training data are not known in advance.

Course Content and Lecturer Information

- **HPC and Tree-Based Methods** (33%) (12 Lectures):

Lecturer: Mr. Stefan Britz Room 5.46 stefan.britz@uct.ac.za

Topics:

1. High Performance Computing
2. Regression and Classification Trees
3. Bagging, Boosting, and Random Forrests.

Prescribed Text: James, G., Witten, D., Hastie, T. and Tibshirani, R., 2013. *An introduction to statistical learning (Vol. 112).* New York: springer.

- **Neural Networks** (33%) (12 Lectures):

Lecturer: Dr. Etienne Pienaar Room 5.43 etienne.pienaar@uct.ac.za

Topics:

1. Perceptrons and the Perceptron Learning Algorithm
2. Neural Networks and Backpropagation
3. Bias-Variance and Regularisation

Prescribed Text: A complete set of course notes with relevant references will be posted on Vula. See also:

Nielsen, M.A., 2015. *Neural networks and deep learning (Vol. 25).* USA: Determination press.

- **Unsupervised Learning** (33%) (12 Lectures):

Lecturer:	Dr. Sebnem Er	Room 5.55 sebnem.er@uct.ac.za

Topics:

1. Clustering
2. Association Rule Mining
3. Self-Organising Maps

Prescribed Text: James, G., Witten, D., Hastie, T. and Tibshirani, R., 2013. *An introduction to statistical learning (Vol. 112).* New York: springer.

- **Course Administrator**: Dr. Etienne Pienaar – etienne.pienaar@uct.ac.za.

Assessment

- **Assignments** (Assignment topics may change):

Assignment 1	Tree-Based Methods	Hazard Score Prediction and Page Block Classification
Assignment 2	Neural Networks	Build a Neural Network and Pattern Recognition
Assignment 3	Supervised Learning	Clustering of Hand-Written Digits

- **Format**: Assignments will consist of a formal write-up to be submitted in PDF format on Vula via the assignments tab. Page limits will be announced upon release of the assignments. Where prediction sets are required, these will be handed in electronically via Vula along with the assignment. Marks will be allocated for the quality of the write up.
- **Course Mark**: The course mark will be constructed from assignments and quizzes/exercises. The deliverables are weighted such that quizzes/exercises count 10% (all equally weighted) of the final mark, the assignments count 40% (all equally weighted) of the final mark.
- **DP Requirement**: A minimum of 35% for the course mark and completion of all assignments and quizzes/exercises by the due dates plus sufficient attendance of lectures.
- **Final Mark**: The final grade for the course will be based on:

Course Mark:	50%.
Final Exam:	50%.
Final Mark:	100%.

Bayesian Analysis

Synopsis

The ideas underlying the Bayesian method dates back to the probability formulations by Thomas Bayes (1763) and Laplace (1812) but have only relatively recently (20-30 years ago) become popular in practical statistical research circles. Much of its popularity is due to the fact that many statistical analyses cannot easily be undertaken using the methods that are usually taught in undergraduate courses at universities but are relatively straight forward to undertake using Bayesian methods.

This module will cover a number of standard computational Bayesian techniques used in statistical analysis. This will include Acceptance-Rejection sampling, Gibbs sampling and Markov Chain Monte Carlo methods. The course is intended to be practical in nature with a large practical component (with a fair amount of rigorous derivations thrown in).

Potential application areas include: finance, ecology, differential equation modelling, biostatistics, spatial modelling, image analysis, state space modelling, variable selection.

Course Content and Lecturer Information

Lecturer: Mr Allan Clark allan.clark@uct.ac.za

(potential) Course Contents

- Introduction to Bayesian analysis.
- Elicitation of prior information.
- Bayesian linear regression.
- Bayesian variable selection and the lasso.
- Markov chain Monte Carlo methods (MCMC):
  - Gibbs sampler, Metropolis Hastings, Data-augmentation, sampling methods.
  - Practical implementation of MCMC (lots of coding in R).
- Mixture modelling.
- Bayesian model averaging.
- Occupancy modelling.

During this course we will cover most of the above list but I am open to including a few other application areas. You should email me suggested topic areas by the end of the first week if you want me to add anything to the course.

Length of module:

The course will be run over a 12 week period. During each week material will be made available at the start of the week. In most cases this will entail reading through lecture slides and/or notes as well as working through accompanying Rmd files. There might be a few videos here and there although this will most probably be kept to a minimum. Weekly catch up sessions will be scheduled to keep track of your progress.

Assessment

- The final grade for the course will be based on the following:

Three small tutorial/practical typed hand-ins	15%
Assignment	40%
Open-book (8 hour) assessment	45%.

- The tutorial/practical assessments should be submitted on Vula by 23 August, 13 September and 8 October respectively. Please diarise these dates. These assessments will be small and will entail understanding the course material covered during the course by the hand-in date. It might entail reading an article and explaining the contents of the paper.
- The assignment will consist of a formal write-up to be submitted in PDF format on Vula via the Assignments tab by 25 October. You will be required to formulate a problem, develop a Bayesian solution and thereafter use R to analyze data. Page limits will be announced upon release of the assignment. Marks will be allocated for the quality of the write up.
- The final assessment will be an open-book take home assessment for you to do in your own time within an 8 hour period. You may consult your notes and any internet sources. You may not confer with anyone and are expected to submit your own work. The assessment will entail answering 'normal exam-type ' questions as well as coding up an analysis of one or two data sets and writing a short report.
- **Missed assessments**: If you have missed an assessment for medical reasons, please send a medical certificate to me immediately. In this case, the course components will be re-weighted appropriately. If you have missed an assessment for any other reason, make sure to also contact me immediately (via email).

BIO-STATISTICS

Lecturers
Prof Francesca Little (Room 6.66) (Convener)
email: Francesca.Little@uct.ac.za
Dr Greg Distiller (Room 6.67)
Email: Greg.Distiller@uct.ac.za

Description
The course material will cover an introduction to the design of medical studies; measures of disease frequency and effect; generalized linear models with specific emphasis on models for binary and count responses; analysis of longitudinal data including GEE and mixed effect models; and survival analysis techniques. This course could be called "Applied generalized (mixed effect) models" and though the applications taught in this course will refer to medical data, the methods have much wider applicability.

Duly Performed Requirements
Timely submission of all class exercises and assignments

Assessment

The course will be assessed through assignments. There will be three assignments

A1: DESIGN assignment contributing 30% towards final mark
A2: GLM assignment contributing 35% towards final mark
A3: LDA+SA assignment contributing 35% towards final mark

A pass mark of 50% is required overall and a sub-minimum of 40% for each assignment.

Lecture times:

Due to the COVID pandemic, we will follow a PDL format. The Biostatistics module will be delivered to students via online resources accessible through VULA. These resources will include a combination of (but not all of) lecture notes, lecture slides, recorded video lectures, class exercises and assignments. Students will be expected to engage with the material through self-study. Lecturers will be available to answer questions on the Q&A forum in VULA and via email. Optional face to face session may be requested and offered by mutual agreement between the lecturers and the students involved.

Students are advised to set aside Tuesday 10 to 12am to start the week's content. They will have to spend more time during the week completing the weekly class exercises. They are welcome and encouraged to communicate with one another when doing the class exercises.

Programming will be done in R but lecture notes may also refer to STATA programs. Student should be able to access STATA via the UCT server, if they wish to program in STATA.

Revised lecture timetable:

Week	Tuesday of week	Planned lectures	Due dates (to be confirmed)
1	15 March	Study design	
2	22 March	Study design	
3	29 March	Measures of Disease Frequency and Effect	Design Assignment hand-out
4	5 April	GLM	
5	12 April	GLM	Design Assignment due
6	19 April	GLM	
7	26 April	Catch-up	GLM assignment hand-out
	3 May	MID-YEAR VACATION	
7	10 May	Longitudinal Analysis 1	GLM assignment due
8	17 May	Longitudinal Analysis 2	
9	24 May	Longitudinal Analysis 3	LDA+SA assignment handout
10	31 May	Survival Analysis	
11	7 June	Survival Analysis	
12	14 June	Survival Analysis	
	21 June	Catch-up	LDA+SA assignment due

Decision Modelling (2020)

Synopsis

The aim of this module is to develop an understanding of human preferences and subjective judgements, for purposes of constructing models for decision support, within which we can apply quantitative tools studied elsewhere. The module thus provides a bridge between technical statistical and operational research tools on the one hand, and interaction with our clients on the other hand (especially in the context of decision support). We study both how to represent human preferences in mathematical models, and what cognitive models and errors may be exhibited by clients when we elicit subjective estimates of underlying parameters.

A substantial portion of the course relates to multiple criteria decision analysis (MCDA), in which we specifically aim to capture preferences and value tradeoffs between conflicting objectives and goals (e.g., social, environmental and economic goals).

In this course, we seek to develop an understanding of the following processes:

- How can we assist decision makers in defining and structuring decision problems in such a way that we are able to apply our quantitative analytical models to them?
- How should we model the preferences and value judgements which are expressed by decision makers or their advisors, in order to generate maximum insight into which courses of action best satisfy these preferences?
- How do people form and express subjective value judgements, for example regarding the relative importance of different goals or of the likelihood of certain events, which we may need to use in our models?

Decision Modelling is a multi-disciplinary field and in addressing the above questions, we touch on issues in psychology, economics, information systems, and operational research.

Course Content and Lecturer Information

- **Part I: Problem Structuring and Approaches to Decision Modelling** (12 lectures):

  Lecturer: Assoc Prof Leanne Scott Room 5.66 Leanne.Scott@uct.ac.za

  Topics: We look at how we can take a general decision problem and structure it in a way that is useful for later modelling. We introduce the three main schools of decision modelling, which offer different ways of modelling how people can and should make multi-criteria decisions. We then focus on one of these schools, value function modelling.

- **Part II: Further Topics in Decision Modelling** (12 lectures):

Lecturer:	Ms Georgina Rakotonirainy Room 5.51
Topics:	We extend the models of the first section into more advanced decision contexts: those involving uncertainty (where measurements cannot be made precisely), preferences over time, or groups of decision makers. As part of this section we also consider problems and biases that people commonly experience when they are asked to think about information or make a judgement.

Assessment

- **Assignments**:

Assignment 1:	Problem structuring of a public sector decision problem (35% of Course Mark)
Assignment 2:	Using a value function to structure, score, weight, aggregate and make an appropriate and defensible decision with regard to a given decision context. (65% of Course Mark)

- **DP Requirement**:A minimum of 35% for the course mark plus attendance of lectures.
- **Exam**: The final grade for the exam will be based on:

Part I:	30%,
Part II:	70%.

- **Final Mark**: The final grade for the course will be based on:

Course Mark:	30%,
Final Exam:	70%.

Introduction to Stochastic Processes (2020)

Synopsis

This module provides a basic overview of stochastic processes with both the state and parameter spaces discrete. Models involving discrete-time stochastic processes are studied and special attention given to stochastic processes known as Markov chains. Two models for the permissible 1-step transitions within the state space will be introduced.

Course Content and Lecturer Information

- **Discrete-valued time series and theory**:

Lecturer:	Mr Dominique Katshunga	Room 5.49	Dominique.Katshunga@uct.ac.za

Topics:

1. Chains and Markov chains
2. Transition probability matrix
3. Time homogeneity
4. Classification of states
5. Probability models for transitions
6. Stationarity and reversibility

Prescribed Text: Lecture notes and slides will be provided as needed.

Assessment

- **Assignments** (Assignment topics may change):

Assignment 1	Possibility & probability models for transition (coding in R will be required)
Assignment 2	Steady-state censuses

- **Format**: Assignments will consist of a formal write-up to be submitted in PDF format on Vula via the assignments tab. Page limits will be announced upon release of the assignments. Marks will be allocated for the quality of the write up.
- **Course Mark**: Each assignment will count 50% toward the course mark.
- **DP Requirement**: A minimum of 35% for the course mark and completion of all assignments by the due dates plus attendance of lectures.
- **Final Mark**: The final exam and course marks will count toward the final mark as follows:

Course Mark:	25%.
Final Exam:	75%.

Introduction to Bayesian Analysis

Synopsis

This module will cover a selection of Bayesian techniques and their applications. The methods and concepts introduced in undergraduate Applied Statistics will be covered in more detail, while new topics will include Gibbs sampling and the Metropolis-Hastings algorithm. While the emphasis will be on practical work (coded in R), students will also be required to understand and derive the mathematical expressions underlying these techniques, where applicable.

Course Content and Lecturer Information

Lecturer:	Patrick Chang	chnpat005@myuct.ac.za
Topics Include:		

1. Introduction and revision.
2. Priors.
3. Gibbs sampling.
4. Linear regression.
5. Metropolis-Hastings.

Format: The course will be delivered in the form of weekly online lectures.

Assessment

- **Assignments and Projects:** One assignment (which counts 40% of final mark) and one project (which counts 60% of the final mark). These assessments will include a combination of derivations and practical applications.
- **Format**: The assignment and project will consist of a formal write-up to be submitted in PDF format on Vula via the assignments tab. Working R scripts will also need to be uploaded. Page limits will be announced upon release of the assessments. Marks will be allocated for the quality of the write up.
- **DP Requirement**: A minimum of 40% for the assignment.
- **Final Mark**: The final grade for the course will be based on:

Assignment:	40%.
Project:	60%.

Likelihood (2020)

Synopsis

This module covers likelihood theory. The likelihood function plays an important role in Bayesian and in frequentist statistics. However, here we will treat likelihood as an approach to inference (point and interval estimates, hypothesis testing, model selection and comparison) in its own right.

In practice, most likelihood methods require substantial computation. Therefore the module will have a strong practical component, and will require the use of R for tutorial questions and the assignment.

Course Content and Lecturer Information

Lecturer: Dr Birgit Erni Room 6.64 birgit.erni@uct.ac.za

Topics:

1. likelihood for different problems, MLE, invariance property, likelihood intervals
2. asymptotic likelihood theory
3. likelihood with multiple parameters, profile likelihood, predictive likelihood, likelihood with random parameters

Prescribed Text: Course notes will be made available on Vula.

Assessment

- The class mark will be based on online quizzes (80%) and participation in online forum discussions (20%).
- The final mark will be based on the class mark (30%) and a project (70%).

Matrix Methods

Synopsis

This course covers basic matrix operations, an introduction to linear algebra, and some pertinent applications of matrix methods in statistics. The mathematical principles covered in this course will be relevant in several other honours modules, especially Multivariate Analysis.

Course Content and Lecturer Information

Lecturer:	Mr. Stefan Britz	Room 5.46	stefan.britz@uct.ac.za

Topics:

1. Nomenclature and revision of basics
2. Determinants
3. Inverses
4. Matrix rank
5. Generalized inverses and linear equations
6. Eigenvalues and eigenvectors
7. Spectral decomposition and singular value decomposition parameters

Prescribed Text: Course notes will be made available on Vula.

Assessment

- **Composition**: One online quiz and one 3-hour invigilated exam.
- **Final Mark**: The final grade for the course will be based on:

Quiz:	10%
Final Exam:	90%

Multivariate Analysis

Synopsis

"Multivariate analysis is the area of statistics that deals with observations made on many variables. The main objective is to study how the variables are related to one another, and how they work in combination to distinguish between the cases on which the observations are made.

The analysis of multivariate data permeates every research discipline: biology, medicine, environmental science, sociology, economics, education, linguistics, archaeology, anthropology, psychology and behavioural science, to name a few, and has even been applied in philosophy. All natural and physical processes are essentially multivariate in nature—the challenge is to understand the process in a multivariate way, where variables are connected and their relationships understood, as opposed to a bunch of univariate processes, i.e. single variables at a time, isolated from one another."
Michael Greenacre, http://multivariatestatistics.org/index.html.

This honours module will introduce students to the underlying theory of multivariate analysis, as well as several multivariate techniques and their application in R.

Lecturer Information

Section	Lecturer	Email
1	Mr Stefan Britz (convener)	stefan.britz@uct.ac.za
2	Prof. Francesca Little	francesca.little@uct.ac.za

Course Content

The syllabus is divided into two broad sections, covering the following topics:

Section 1: Theory and Inference

1. Introduction to and Visualisation of Multivariate Data
2. Singular Value Decomposition, Eigenvalue Decomposition and Spectral Decomposition Revisited
3. The Multivariate Normal Distribution
4. Multivariate Maximum Likelihood Estimation
5. Multivariate Inference
6. MANOVA and Multivariate Regression Analysis

Section 2: Applications of Multivariate Analysis

1. Principal Component Analysis and Factor Analysis
2. Canonical Correlation Analysis
3. Discriminant Analysis
4. Correspondence Analysis

Prescribed Textbook

RA Johnson & DW Wichern, "Applied Multivariate Statistical Analysis", 6th edition, Pearson International Edition, 2007.

Additional Resources:

B Everitt, "An R and S-Plus Companion to Multivariate Analysis:", Springer, 2005.
AC Rencher, "Methods of Multivariate Analysis", second edition, Wiley, 2002.

Assessment

Assignments

Each section will be followed by an assignment.

- Assignment 1 – Theory and Inference
- Assignment 2 – Application of Various Multivariate Techniques

Assignments will consist of a formal write-up to be submitted in PDF format on Vula. Page limits and time for completion will be announced upon release of the assignments. Assignments may involve derivation, calculation, computing and interpretation of results obtained from the analysis of data sets.

Quizzes

There will be a weekly quiz testing basic concepts and application of that week's work. These will be submitted via Vula and will be due at 09:00 on the following Monday.

Final Mark: The final grade for the course will be based on:

Quizzes (best 10/12):	10%
Assignment 1:	45%
Assignment 2:	45%

A final mark of 50% is required to pass. Also, both assignments and at least 10 out of the 12 quizzes must be submitted (what constitutes a "submission" is at the discretion of the convener).

Operations Research A

Synopsis

Operations Research (OR) is the study of scientific approaches to decision-making. Through mathematical modelling, OR seeks to design, improve and operate complex systems in the best possible way. OR has many applications in science, engineering, economics, and industry and thus the ability to formulate and solve OR problems is crucial for both researchers and practitioners. The mathematical tools used for the solution of such models are either deterministic or stochastic, depending on the nature of the system modelled. The goal of this course is to teach you to formulate, analyse, and solve mathematical models that represent real-world problems. The course will also introduce you to how to use Open Solver an open source add in to Excel for solving optimization problems expressed as spreadsheet models. Excel will also be used for simulation modelling. Note: This module is intended for students who have not completed STA3036S

Course Content and Lecturer Information

Lecturer:

- Dr Juwa Nyirenda, Room 6.68, Juwa.Nyirenda@uct.ac.za
- Dr Georgina Rakotonirainy, Room 5.51, rosephine.rakotonirainy@uct.ac.za

Topics:

1. Linear Programming formulation and Solution
2. Optimality Principles
3. Further Practicalities
4. Network and Project management models
5. Queuing Theory
6. Simulation

Prescribed Text: Complete notes for the course will be provided.

Assessment:

- **Composition:**
  - ***Queuing Theor***y: One assignment + Simulation Project
  - ***Linear Programming***: One assignment + 6 Tutorial questions (quizzes)
- **Format**: Assignments will consist of a formal write-up to be submitted in PDF format on Vula via the assignments tab. Page limits will be announced upon release of the assignments. Marks will be allocated for the quality of the write up.

Course Mark:

- Each assignment counts 25% of the course mark; the simulation project counts 25% of the course mark and average of 6 tutorial questions (quizzes) will count 25% of the course mark.
- **DP Requirement**: A minimum of 35% for the course mark and completion of all assignments by the due dates.
- **Final Mark:** The course grade.

Operations Research B

Synopsis

Operations Research (OR) is the study of scientific approaches to decision-making. Through mathematical modelling, OR seeks to design, improve and operate complex systems in the best possible way. OR has many applications in science, engineering, economics, and industry and thus the ability to formulate and solve OR problems is crucial for both researchers and practitioners. The mathematical tools used for the solution of such models are either deterministic or stochastic, depending on the nature of the system modelled. The goal of this course is to teach you to formulate, analyse, and solve mathematical models that represent real-world problems. The course will also introduce you to how to use Open Solver an open source addin to Excel for solving optimization problems expressed as spreadsheet models. For simulation modelling a dedicated simulation package will be used.

Note: This module is intended for students who have completed STA3036S or Operations Research A.

Course Content and Lecturer Information

- **Lecturers:**

Dr Juwa Nyirenda Room 6.68 Juwa.Nyirenda@uct.ac.za

Dr Georgina Rakotonirainy Room 5.51 rosephine.rakotonirainy@uct.ac.za

- **Topics:**

1. Multi-objective and Goal Programming.
2. Duality.
3. Data Envelopment Analysis.
4. Integer Programming.
5. Deterministic and Stochastic Dynamic programming.
6. Simulation.

- **Prescribed Text:**

Complete notes for the course will be provided.

- **Assessment**

Topics 1,2,3 & 4: One assignment (20%) + Average of 6 tutorial questions (quizzes) (30%).

Topic 5: Two assignments (15% each).

Topic 6: One simulation Project (20%).

- **Format:**

Assignments will consist of a formal write-up to be submitted in PDF format on Vula via the assignments tab. Page limits will be announced upon release of the assignments. Marks will be allocated for the quality of the write up.

- **Course Mark:**

The three assignments will count 50% of the course mark; the simulation project counts 20% of the course mark and the average of 6 tutorial questions (quizzes) will count 30% of the course mark.

- **DP Requirement:**

A minimum of 35% for the course mark and completion of all assignments by the due dates.

- **Final Mark:**

The final grade = Course grade

Portfolio Theory (2020)

Synopsis

The course aims to familiarise the student with the foundations of portfolio theory, specifi cally as they related to the quadratic asset allocation problem driven by daily equity, bond and money market returns. The course will use R to model tactical and strategic assets via industry sector data sourced from INET. The course aims to motivate the use of equilibrium models in finance and use the link between asset pricing, financial equilibrium and mean-variance efficiency to make data informed portfolio selection decisions. The course content on the vula site will include extra lecture notes that are not part of the curriculum and a variety of additional readings that extend the content for those that have the time and interest. The course takes an ergodicity economics perspective of finance within the traditional mean-variance framework.

Note: This course forms the pre-requisite course for STA5091Z.

Course Content and Lecturer Information

Lecturer: A/Prof Tim Gebbie Room 5.39 tim.gebbie@uct.ac.za

Topics:

Week 1: From Gold to Markets

Week 2: Data Wrangling in R

Week 3: Static Mean-Variance Theory

Week 4: Tactical Asset Allocation

Week 5: Black-Litterman Model

Week 6: Fundamental Law of Active Management

Week 7: Insights into Active Management

Week 8: The Roll Criticism

Week 9: Estimation Uncertainty

Week 10: Performance Measures

Week 11: Performance Attribution

Week 12: Indexation

Readings: See the below chapters, and aims to summarise key ideas in no more than 12, 2 hour sessions, at the level of the first year of graduate school (first year science postgraduate) :

1. Elton, E., Gruber, M., Brown, S., and Goetzmann, W.,(2014), Modern Portfolio Theory and Investment Analysis, 9-th edition, Wiley [Chapters 13, 17, 26, 28]

2. Ingersoll, J. E., (1987), Theory of Financial Decision Making, [Chapter 11, 13]

3. Lee, W., (2000), Theory and Methodology of Tactical Asset Allocation [Chapters 2, 3, 7]

4. Prigent, P-L., (2007), Portfolio Optimization and Performance analysis. Chapman & Hall [Chapters 1,3,4,5,7]

5. Rachev, S. T., Hsu, J. S. J., Bagasheva, B. S., Fabozzi, F. J., (2008), Bayesian Methods in Finance, [Chapters 8, 13, 14]

6. Ross, S., (2005), Neoclassical Finance, Princeton Press [Chapters 1]

Assessment

I will split the course grades: 40% to the 2 assignments, 25% to the online theory essay and 25% for the cap-stone assignment project.

- **Composition:** Two assignments (theory and code) make up the course mark, there will be an online theory assessment and a cap-stone practical take-home/online assessment. Course notes and meet-up times are found from the Vula Lessons tabs which align with the weekly content schedule.

- **Format**: Assignments will consist of a formal write-up to be submitted in PDF format on Vula via the assignments tab. R markdown and Latex templates are available from the course Resources on the Vula site. The capstone assessment will require code based submission in R or R markdown within a timed 12-24 hour period.

- **Course Mark**: In calculating the course mark, each assignment counts at least one third of the **course mark**. There be at least one online presentation session where results for the specific assignments will be discussed, these presentations will then inform the class mark, and will be student presentations given using beamer or equivalent in Latex prepared PDF documents. The course grade will be 50%.

- **DP Requirement**: The course has an exam subminimum of 45% on the combination of timed online cap-stone take-home practical assessment and the timed online theory MCQ. The submission of all assignments by the due dates is expected. Online lecture and meet-up attendance is compulsory and active engagement with class discussions is expected.

- **Final Mark**: The final grade for the course will be based on:

Course Mark:	40% (2 Assignments.)
Final Assessments:	60% (25% Theory Essay Questions + 25% Practical Take-home Exam.)

Statistical Computing (2020)

Synopsis

Statistical Computing refers to the use of computers to do statistics. This definition includes data handling, visualization, statistical modelling, and writing software for statistical analyses. Statistical computing is also used to refer to computer intensive methods where the computer is used to solve problems that are analytically difficult (then also known as *Computational Statistics*). This latter category includes Monte Carlo methods and bootstrapping.

The R language for statistical computing will be used in this course. We start with the basics: R data structures such as vectors, matrices, and data frames. You will learn how to manipulate, summarize and organize data and data files, use R to create quality graphics, program and write your own functions, use simulations to solve statistical problems (Monte Carlo and bootstrap), optimize functions, and finally how to do parallel computing in R.

Course Content and Lecturer Information

Lecturer: Dr Birgit Erni Room 6.64 birgit.erni@uct.ac.za

Topics:

1. data structures
2. programming basics
3. data management (summarizing, transforming between formats, aggregating)
4. graphics
5. R Markdown for reproducible research
6. Monte Carlo methods
7. bootstrap
8. optimization
9. parallel computing

Prescribed Text: Course notes for the sections on Monte Carlo, bootstrap and optimization will be made available on Vula. A list with resources for learning R will be given.

Assessment

For 2020, a *project* will replace the final practical exam. The final mark will be calculated as:

Test early during the semester	20%
Project	80%

Time Series (2020)

Synopsis

Many things that we measure or collect data on, change over time. Data points correlated in time are common in many fields: economics, environment, medicine, engineering. In recent years, new types of data recording devices and data sources have resulted in increased amounts of large time series data sets. Also, more powerful computers allow for more powerful statistical methods to be used. This course starts off with a brief revision of the Box-Jenkins approach to time series modelling. Then we will explore various methods in modern time series analysis, including spectral analysis, filtering, state-space models, Kalman filters.

Course Content and Lecturer Information

Lecturer:	Sulaiman Salau	Room 5.54	sulaiman.salau@uct.ac.za
	Birgit Erni	Room 6.64	birgit.erni@uct.ac.za

Topics:

1. spectral analysis and filtering
2. state-space models
3. further topics in modern time series analysis, should time allow

Prescribed Text: Robert H. Shumway & David S. Stoffer, 2016. *Time Series Analysis and its Applications. With R Examples.* 4th edition. Springer. https://www.stat.pitt.edu/stoffer/tsa4/

Assessment

There will be two assignments. Up to 20% of the course mark could be based on class participation. The final grade for the course will be calculated as:

Course Mark:	40%.
Final Exam:	60%.