

Faculty of Natural and Agricultural Sciences

Fakulteit Natuur- en Landbouwetenskappe Lefapha la Disaense tša Tlhago le Temo

Department of Statistics Postgraduate programmes

There are plenty of opportunities ahead. This is really a GOLDEN age for Statistics

Sara van den Geer- President of the Bernoulli Society (May 2017)

Last revision: September 2023

Pathways



The Department of Statistics at UP, is the only Statistics Department in the country with the following unique research vehicles

South Africa DST-NRF-SAMRC SARChI Research Chair in Biostatistics

This research chair in Biostatistics and its research group will focus on solving inter-disciplinary problems commonly seen in Africa through research and capacity development in biostatistical analytics. This research group will develop computational software and tools to publicise these novel methodologies and to facilitate their applications and implementations in public health research. Furthermore, this research group will train future research leaders and qualified biostatistics researchers through a series of short courses and symposia in biostatistics.

Funding

Competitive postgraduate bursaries requiring self-application are made available by stakeholders such as SARChI and Statomet (both situated within the Department of Statistics at UP), independent NRF freestanding scholarships, South African Statistical Association honours bursaries, and the Centre for Artificial Intelligence Research (CAIR). Excellent academic achievement is used as a benchmark for bursary awards. Priority for bursaries are given to South African Citizens and Permanent Residents. The Department of Statistics at the University of Pretoria offers an extensive and flexible range of postgraduate programmes.

The following postgraduate programmes and disciplines are presented (in collaboration with other departments at UP and other institutions):



NOTE: The BSc (Hons) Statistics and Data Science degree can only admit students completing STK undergraduate modules in programmes <u>not</u> within the EMS faculty i.e. without a BCom undergraduate degree (for example, BSocSci or similar undergraduate degrees).

Honours Programmes

General

The honours degree is presented in Mathematical Statistics and in Statistics and Data Science. Apart from the modules presented by the Department of Statistics, it is also possible to include postgraduate modules from other departments e.g. Mathematics, Economics, etc. Regarding the number of credits from other departments the following is applicable:

- a. The department must be a "related" department.
- b. The student must hand in a proposal to the head of the department for approval at their own discretion.
- c. At most <u>one</u> of the six modules (required for the degree) can be replaced with the understanding that compulsory modules cannot be replaced and thus only elective modules are replaceable.
- d. If considered/granted, the replacement module must be of the same number of credits as other coursework modules in the Department of Statistics (i.e. 15 credits exactly).

Normally a module is only offered once a year in any one semester, see program composition for presentation period (as well as timetable for that semester/year, as available on the departmental website). A candidate must compile their programme in consultation with the head of the department or the department's postgraduate committee.

Admission requirements

- a. A relevant Bachelor's degree with Mathematical Statistics or Statistics on the 300 level is required (Economics on the 300 level is also required for BCom(Hons) in Statistics and Data Science specialising in Economic Statistics).
- b. For BSc(Hons) and BCom(Hons) in Mathematical Statistics, an average mark of 65% or more for Mathematical Statistics modules on the 300 level. Admission can also be considered for students from other programmes with at least the modules WST 311, 312, 321 as the second major without STK 353, if it is academically warranted.
- c. For BCom(Hons) in Statistics and Data Science an average mark of at least 65% in Statistics 310, 320 and 353 is required. Admission can also be considered for students from other BCom programmes with at least the modules STK 310, 320 as the second major without STK 353, if it is academically warranted.
- d. Please see the additional requirements for external students as well as international students at the end of this document. Student numbers are limited to a maximum of 40, collectively over all Honours Programs in the Department of Statistics.
- e. Historical performance during prior studies will also be considered in selecting students. Specific attention will also be given to modules repeated and duration of study.
- f. Any additional entrance requirements as specified by the head of the department in consultation with the departmental postgraduate selection committee. Students with time gaps in between studies may be required to write an entrance evaluation.
- g. A student may only commence studies in this department at the beginning of an academic year (January/February).

Duration of programme

The minimum duration is one year of full-time study or two years of part-time study. A student must complete their study for an honours degree, in the case of full-time students, within two years from the first examination to the final examination and in the case of part-time students, within three years from the first examination to the final examination. Under special circumstances, the Dean, on the recommendation of the head of department, may approve a limited extension of this period.

Promotion

The progress of all honours candidates is monitored biannually by the department's postgraduate committee and head of department. Subject to exceptions approved by the dean, on recommendation of the head of the department, a student may not enter for the honours' examination in the same subject more than twice and a candidate's study may be terminated if the progress is subsequently deemed unsatisfactory or if the candidate is unable to finish their studies during the prescribed minimum period. The research report may only be registered for once – unsuccessful completion will result in dismissal from the degree.

The following applies to Honours and Masters study in the Department of Statistics:

1) A student will be allowed the opportunity to write a special exam in January of the following year on the condition that the student has only failed one module AND the failure mark is between 40 and 49 percent, inclusive. The student SHOULD have passed all the other modules.

2) Paragraph 1) above is subject to the following limitations:

a) Honours degree: None of the modules will be repeated more than once. The exception is that modules Research Report WST/STK 795 cannot be repeated, and any failure in either of these two modules will result in the exclusion from the honours degree. No special exams for Research Orientation STK796.

b) Masters degree: None of the modules will be repeated more than once. No special exams for Research Orientation STK899.

3) A student who fails one or more modules will be excluded. In this case, the student has the opportunity to lodge an appeal with the relevant faculty.

4) Sick exams will be scheduled for January of the following year. All sick cases must be approved by the Head of the Department of Statistics in writing. No special exams are made available after a sick exam.

Programme composition

The available programs with their compositions are given below. Students having majored in Mathematical Statistics (WST) in a BSc qualify only for the BSc(Hons) Mathematical Statistics. Students having majored in Mathematical Statistics (WST) in a BCom undergraduate degree qualify only for the BCom(Hons) Mathematical Statistics. Students having majored in Statistics (STK) in a BSc (or other non-BCom degree) qualify for BSc(Hons) Statistics and Data Science. Students having majored in Statistics (STK) in a BCom qualify for the BCom(Hons) Statistics and Data Science.

BSc(Hons) in Mathematical Statistics and BCom(Hons) in Mathematical Statistics (135 credits):

First Semester	Second Semester	
Compulsory modules (60 credits)		
LMO710 Linear Models (15)		
MVA710 Multivariate Analysis (15)		
WST795 Research Report: Mathematical Statistics (30)		
STK796: Research Orientation (0)		
Elective modules (choose 5) (75 credits)		
SFT720 Sampling Techniques (15)	LMO720 Linear Models (15)	
VMT710 Distribution-free Methods (15)	MVA720 Multivariate Analysis (15)	
STC 710 Simulation and computation (15)	PNP720 Parametric Stochastic Processes (15)	
EKT 725 Text and behavioural analytics (15)	EKT720 Introduction to Statistical Learning (15)	
	STK781 Linear Mixed Models (15)	
	STC 720 Capita selecta: statistics 720 (15)	

BSc(Hons) in Statistics and Data Science (135 credits):

First Semester	Second Semester	
Compulsory modules (60 credits)		
RAL780 Regression Analysis (15)	MET720 Multivariate Techniques (15)	
	EKT720 Introduction to Statistical Learning (15)	
STK795 Research Report: Statistics (30)		
STK796: Research Orientation (0)		
Elective modules (choose 4) (75 credits)		
SFT720 Sampling Techniques (15)	FIN 707 (30) (year module)*	
EKT 725 Text and behavioural analytics (15)	STK781 Linear Mixed Models (15)	
STC 710 Simulation and computation (15)	STC 720 Capita selecta: statistics 720 (15)	

*FIN 707 can only be taken on approval by the Head of Department

First Semester	Second Semester	
Compulsory modules (75 credits)		
RAL780 Regression Analysis (15)	EKT720 Introduction to Statistical Learning (15)	
	MET720 Multivariate Techniques (15)	
STK795 Research Report: Statistics (30)		
STK796: Research Orientation (0)		
Elective modules (choose 3) (45 credits)		
SFT720 Sampling Techniques (15)	STK781 Linear Mixed Models (15)	
MIE780 Microeconomics (15)	FIN 707 (30) (year module)*	
EKT 725 Text and behavioural analytics (15)	STC 720 Capita selecta: statistics 720 (15)	
MEK780 Macroeconomics (15)		
STC 710 Simulation and computation (15)		

• BCom(Hons) in Statistics and Data Science (120 credits):

*FIN 707 can only be taken on approval by the Head of Department

Notes

- a. Some of the second semester modules have first semester modules as prerequisites, as explicated in the section "Particulars regarding the contents of modules."
- b. Tests and assignments may be taken into account in the determination of the final mark in each module. Exam entrance may be enforced in each module <u>on discretion of the lecturer</u>.
- c. General Regulation G.17 applies with regard to examination registration.
- d. Examinations in a module will take place during one of the official university examination time periods. In order to pass the programme, a final mark of at least 50% in each of the prescribed modules must be obtained. The honours examination can be divided into at most four sections which must each be written in a time period of at most three weeks and over at most four semesters.
- e. It is compulsory for all the postgraduate students to comply with requirements for the module STK796 Research Orientation.
- f. Study financing remains the responsibility of the student. The department is not responsible for the funding or obtaining funding for a student. A student must thus plan their finances personally. The department will assist in directing students towards bursary opportunities.

In the case of a dispute with the allocated study leader(s) of research report, the department will proceed as follows:

- The student must raise their valid concerns in an open professional manner, without fear of retribution.
- The supervisor(s) has to find constructive solutions to the problem and obtain agreement from the student for the way forward.
- The above-mentioned process assumes that the student did not commit an act which involves disciplinary actions, such cases should be referred to the legal department.
- If the conflict/dispute cannot be solved amicably between student and supervisor then either party can request a change in supervisor. If the current supervisor agrees **and** a suitable 'new' supervisor can be found then the student may continue with their existing studies. Allocation of a new supervisor is subject to capacity constraints of the department and the current research topic cannot be guaranteed. All existing financial commitments shall be honoured.
- Finally, in absence of the above-mentioned agreement the student must terminate their current studies.
- g. A student should be available for exams for the entire official exam session in both June and November.

Research Report: WST 795 and STK 795

- a. The total time spent on the essay must be at least double that of any module. The submission date of the research report is around mid-October each year. Any additional dates/requirements for the successful completion of WST/STK 795 will be communicated in that year.
- b. Refer to the document: Honours Research Report Course Information available on the web (<u>www.up.ac.za/statistics</u>) under Postgraduate Studies. Registration and allocation of topic (and by implication a study leader) will be dealt with by the Head of the Department and research report coordinators during February each year. Details will be conveyed by email

before the start of the academic year and a number of topics will be provided on clickUP of WST/STK 795 prior to allocation for consultation.

- c. There is a preparation course prior to the commencement of the academic year comprising part of the compulsory module STK796. *This course will take place during working hours so students should organise leave in advance.* These classes are compulsory. Details in the document mentioned above.
- d. Lectures commence on in the academic week that UP undergraduate modules commence (modules from the Economics Department may start earlier please check their dates).
- e. This module may only be done once. If a student fails WST 795/STK 795 the student will be excluded from the honours program with no option to appeal.
- f. Part-time students are advised to do the module WST 795/STK 795 during their second year of study and complete most modules during the first year of study. STK 796 has be completed in both years of study.
- g. The research report and any additional expected submissions must be handed in together with an acceptable TurnItIn report.

Particulars regarding the contents of modules (see the <u>yearbook</u> for detail as well)

Distribution-free Methods 710 (VMT 710)

Prescribed book:

Nonparametric Statistical Inference by Gibbons and Chakraborti, Taylor and Francis Group.

Curriculum:

A selection of: Nonparametric stochastic processes. Power and asymptotic power of distribution-free procedures. Theory and simulation. Asymptotic relative efficiency. Linear rank tests: Definition, properties and applications. Equal in distribution technique. Counting and ranking statistics. Introduction to one and two sample *U*-statistics. Permutation and distribution-free rank-like statistics. Multi-sample distribution-free tests, rank correlation and regression. Some nonparametric bootstrap and smoothing methods. Prerequisite: Admission to either BSc(Hons) or BCom(Hons) Mathematical Statistics

Introduction to Statistical Learning 720 (EKT 720)

Prescribed book:

- Gujarati, D.N.: Basic Econometrics (Second edition) (McGraw-Hill, 1988)
- Intriligator, M.D.; Bodkin7 G.C. and Hsiao, C.: Econometric Models, Techniques & Applications (Prentice Hall, Inc., 1996)
- Hastie, T; Tibshirani, R and Friedman, J: The Elements of Statistical learning (Second edition) (Springer 2009)

Recommended books:

- Pindyck, R. S. and Rubinfield, D.L.: Econometric Models and Economic Forecasts (Fourth edition) (McGraw-Hill, 1991)
- > Draper, N and Smith, H: *Applied Regression analysis* (John Wiley, 1998)

Curriculum:

The emphasis is on the theoretical understanding and practical application of advanced statistical modelling. The following topics are covered:

Single equation models: Nonparametric regression. Bootstrap procedures within regression analysis, k-nearest neighbour classification. Modelling categorical response variables. Penalised regression and feature selection. Non-linear regression modelling. Factor analysis regression. Some new developments in regression and classification.

Simultaneous equation models: Specification, identification and estimation of simultaneous equation models. <u>Prerequisites:</u> Entrance to the degree and for BCom(Hons) Statistics: RAL 780

Linear Mixed Models 781 (STK781)

Specification of linear mixed model, model assumptions, estimation (REML and ML), diagnostics, hypothesis tests, interpretation of parameter estimates, calculating predicted values. Specific models: two- and three-level models for clustered data, intraclass correlation coefficients, repeated measures data, random coefficient models for longitudinal data, models for clustered longitudinal data, models for data <u>Prerequisites:</u> Entrance to the degree

- Linear Models 710 (LMO 710)

Recommended book:

Searle, S.R.: *Linear Models* (John Wiley, 1971)

Curriculum:

Projection matrices and sums of squares of linear sets. Estimation and the Gauss-Markov theorem. Generalized t and F tests.

Prerequisite: Admission to either BSc(Hons) or BCom(Hons) Mathematical Statistics

Linear Models 720 (LMO 720)

Recommended book:

Lecture notes

Curriculum:

The singular normal distribution. Distributions of quadratic forms. The general linear model. Multiple comparison. Analysis of covariance. Generalized linear models. Analysis of categorical data.

Prerequisite: [LMO 710]

- Macroeconomics (MEK 780) Refer to the Department of Economics for details.
- Microeconomics (MIE 780)

Refer to the Department of Economics for details.

- Multivariate Analysis 710 (MVA 710)

Recommended books:

Johnson, R.A. and Wichern, D.W.: Applied Multivariate Statistical Analysis (Sixth edition) (Prentice Hall)

➢ Morrison, Donald, F.: Multivariate Statistical Methods. (Fourth edition)(Thomson) <u>Curriculum:</u> Matrix algebra. Some multivariate measures. Visualizing multivariate data. Multivariate distributions. Samples from multivariate normal populations. The Wishart distribution. Hotelling's T^2 statistic. Inferences about mean vectors. Related current modern topics.

Prerequisites: Admission to either BSc(Hons) or BCom(Hons) Mathematical Statistics

Multivariate Analysis 720 (MVA 720)

Recommended books:

> Johnson, R.A. and Wichern, D.W.: *Applied Multivariate Statistical Analysis* (Sixth edition) (Prentice Hall).

Everitt, B. and Torsten Hothorn, T. An introduction to applied multivariate analysis with R

<u>Curriculum:</u>

The matrix normal distribution, correlation structures and inference of covariance matrices. Discriminant analysis. Principal component analysis. The biplot. Multidimensional scaling. Exploratory factor analysis. Confirmatory Factor analysis and structural equation models. Related current modern topics Prerequisite: [MVA 710]

Multivariate Techniques 720 (MET 720)

Prescribed book:

Bain, L.J. and Engelhardt, M.: Introduction to Probability and Mathematical Statistics (Second Edition) (PWS-Kent Publishing Company, 1992)

<u>Curriculum</u>

Point and Interval estimation. Sampling distributions, central limit theorem, simulations and Bootstrap. Bayesian inference, posterior distribution, Hypotheses testing using confidence intervals, ratio tests, simulated null distributions and power function <u>Prerequisites:</u> Admission to either BCom(Hons) or BSc(Hons) Statistics and Data Science.

Parametric Stochastic Processes 720 (PNP 720)

Recommended book:

Baht, U.N: An Introduction to Queueing Theory: Modelling and Analysis in Applications (Birkhäuser, 2008).

<u>Curriculum:</u>

Introduction to statistical measure theory. Queueing processes: M/M/1; M/M/S; M/G/1 queues and variants; limiting distribution of the queue length and waiting times. Queueing networks. Some stochastic inventory and storage processes.

Prerequisite: Admission to either BSc(Hons) or BCom(Hons) Mathematical Statistics

Regression Analysis 780 (RAL 780)

Prescribed book:

Kutner, Nachtsheim and Neter (2004). Applied Linear Regression Models, 4th edition, Mc Graw Hill

Recommended books:

Freund, J.F., Wilson, J.W. and Ping, S. (2006). Regression Analysis: Statistical Modeling of a Response Variable, 2nd edition, Elsevier.

- Freund, J.F. and Ramon, C.L. SAS System for Regression, 3rd edition, Cary, N.C. : SAS Institute.
- Mendenhall, W. and Sincich, T.: A second module in business statistics: Regression Analysis 6th edition, Macmillan
- Montgomery, D.C. and Peck, E.A.: Introduction to Linear Regression Analysis, 2nd edition, John Wiley.

Curriculum:

Matrix methods in statistics. Simple and multiple regression models. Sums of squares of linear sets. Generalized t- and F-tests. Residual analysis. Diagnostics for leverage, influence and multicollinearity. Indicator variables. Regression approach to analysis of variance. Weighted least squares. Ridge regression. Theory is combined with practical work.

Prerequisites: Admission to either BCom(Hons) or BSc(Hons) Statistics and Data Science.

Sampling Techniques 720 (SFT 720)

Prescribed book:

Lohr, S.L.: Sampling: Design and Analysis (2nd edition) (Brooks/Cole, Cengage Learning, 2010)

<u>Curriculum:</u>

Simple random sampling. Estimation of proportions and sample sizes. Stratified random sampling. Ratio and regression estimators. Systematic and cluster sampling. Introduction to spatial statistics. Spatial sampling – both model and design based approaches. **Prerequisites:** Admission to either BSc(Hons) or BCom(Hons) Mathematical Statistics or BCom(Hons) or BSc(Hons) Statistics and Data Science

Simulation and computation 710 (STC 710)

<u>Curriculum:</u>

Efficient programming, Monte Carlo simulation, sampling of discrete and continuous probability models, General transformation methods, Accept-reject methods, Monte Carlo integration, importance sampling, numerical optimisation, Metropolis-Hastings algorithm, GIBBS sampling.

<u>Prerequisites:</u> Admission to either BSc(Hons) or BCom(Hons) Mathematical Statistics or BCom(Hons) or BSc(Hons) Statistics and Data Science

Text and behavioural analytics 725 (EKT 725)

<u>Curriculum</u>: Clustering and clustering extensions. Mixtures of distributions and regressions. Latent components, soft allocation and belongings. Applications in unstructured data, including text data. Identification and interpretation of behavioural patterns.

<u>Prerequisites:</u> Admission to either BSc(Hons) or BCom(Hons) Mathematical Statistics or BCom(Hons) or BSc(Hons) Statistics and Data Science

Capita selecta: Statistics 720 (STC 720)

<u>Curriculum</u>: This module considers specific topics from the diverse field of statistics as deemed supportive towards the training of the cohort of scholars.

<u>Prerequisites:</u> Admission to either BSc(Hons) or BCom(Hons) Mathematical Statistics or BCom(Hons) or BSc(Hons) Statistics and Data Science

Statistical Process Control 780 (SPC 780) Currently not presented. <u>Prescribed book:</u>

 Montgomery, D.C.: Statistical Quality Control: A Modern Introduction (Sixth Edition - International Student Version) (John Wiley, 2009)

<u>Curriculum:</u>

Quality control and improvement. Shewhart, cumulative sum (CUSUM), exponentially weighted moving average (EWMA) and Q control charts. Determining process and measurement systems capability. Parametric and nonparametric (distribution-free) control charts. Constructing control charts using Microsoft Excel and/or SAS. Obtaining run-length characteristics via simulations, the integral equation approach, other approximate methods and the Markov-chain approach.

<u>Prerequisites:</u> Admission to either BSc(Hons) or BCom(Hons) Mathematical Statistics or BCom(Hons) or BSc(Hons) Statistics and Data Science

Research Orientation 796 (STK 796)

Curriculum:

The yearbook indicates that to comply with requirements for this module, students need to

- 1. Attend the research orientation bootcamp (which was hosted in February);
- 2. Attend 3 departmental/other approved seminars; and
- 3. Be registered for the research orientation module each year of registration for the postgraduate degree in question.

Regarding point 2 above:

Students will be expected from 2023 onwards (from now) to attend at least 3 departmental seminars (in person/online where available) and submit a 2-3-page report in October each year of registration via clickUP with a brief summary of each seminar with regards to the topic, speaker, and aspects of the topic presented at the seminar.

Submissions will then be taken into consideration to recommend that a student complies with requirements for the research orientation module.

Note that:

Even if you only complete coursework this year and complete your research report in the following year, you still need to be registered and comply with requirements for the research orientation module in each year of registration.

If you do not comply with requirements for the research orientation module, you will be unable to graduate as the research orientation module forms part of the yearbook description of the degree that you are enrolled for.

If you have already attended and complied with requirements for attendance of the research orientation bootcamp (for example, in your honours year) you need not attend it again in a subsequent year/for a subsequent postgraduate degree in this department.

If you attend and present at a conference/symposium on work regarding your research component, you can include this as part of the summary.

Master's Programmes

General

The Master's degree is presented in the following directions: Advanced Data Analytics (Mathematical Statistics), Advanced Data Analytics (Statistics), eScience.

Students must consult their relevant faculty to ascertain themselves with the applicable submission dates for the dissertation, essay or thesis. Students must frequent themselves with the regulations regarding the submission of dissertations, essays and theses in order to comply with the requirements of their host faculty.

The following degrees are offered:

- MSc in Advanced Data Analytics
- MCom in Advanced Data Analytics
- MSc in eScience

Admission Requirements

a. A relevant honours degree as indicated is required:

Programme:	Prerequisite:
MSc in Advanced Data Analytics	BSc(Hons) in Mathematical Statistics
	Or
	BSc(Hons) Statistics and Data Science
MCom in Advanced Data Analytics	BCom(Hons) in Statistics
	Or
	BCom(Hons) in Mathematical Statistics
MSc in eScience	An honours degree in either statistics, mathematics, computer science, physics, or related fields; AND demonstrable knowledge of basic principles of probability and statistics, computing, calculus and linear algebra; OR
	Passing an entrance evaluation designed by the academic advisory committee of the programme

- a. For MSc (Advanced Data Analytics) an average mark of at least 65% for coursework in BSc(Hons) Mathematical Statistics or BSc(Hons) Statistics and Data Science modules, with a minimum of 65% for the research component.
- b. For MCom (Advanced Data Analytics) an average mark of at least 65% for coursework in the respective BCom(Hons), with a minimum of 65% for the research component.
- c. For MSc eScience an average of 65% at honours level is the minimum for consideration, although admission will be competitive and an honours average of at least 70% is highly recommended. Student numbers are limited to 6 at UP. Admission is additionally dependent on availability of supervisor/s and/or projects within the participating departments.

- d. Please see the additional requirements for external students as well as international students at the end of this document. Student numbers are limited to a maximum of 26 collectively over all Master's Programs in the Department of Statistics.
- e. Historical performance during prior studies will also be considered in selecting students. Specific attention will also be given to modules repeated and duration of study.
- f. Admission to MSc (Advanced Data Analytics) and MCom (Advanced Data Analytics) is also subject to the availability of supervisory capacity in the field of specialisations available in the department.
- g. The research proposal of applicants should be in line with the research focus of the department or participating departments in the case of MSc eScience.
- h. Any further additional entrance requirements as specified by the head of the department in consultation with the department's postgraduate selection committee. Students with time gaps in between studies may be required to write an entrance evaluation.
- i. The Head of Department, in consultation with the departmental postgraduate selection committee and participating departments (in the case of MSc eScience) reserves the right to prescribe additional modules; understanding that additional modules need to be successfully completed before continuation in the program, and may be for non-degree purposes.
- j. A student may only commence studies in this department at the beginning of an academic year.

Duration

As long as progress is satisfactory, renewal of registration of a master's student will be accepted for a second year of study in the case of a full-time student. Renewal of registration for a third and subsequent years for a full-time student will only take place when Student Administration of the Faculty receives a written motivation (including a timeline (Gantt chart) indicating potential completion date) that is supported by the Head of Department and Postgraduate Committee.

Promotion

The progress of all master's candidates is monitored biannually by the supervisor and the department's postgraduate committee. A candidate's study will be terminated if the progress is unsatisfactory or if the candidate is unable to finish their studies during the prescribed period.

The following applies to Honours and Masters study in the Department of Statistics:

1) A student will be allowed the opportunity to write a special exam in January of the following year on the condition that the student has only failed one module AND the failure mark is between 40 and 49 percent, inclusive. The student SHOULD have passed all the other modules.

2) Paragraph 1) above is subject to the following limitations:

a) Honours degree: None of the modules will be repeated more than once. The exception is that modules Research Report WST/STK 795 cannot be repeated, and any failure in either of these two modules will result in the exclusion from the honours degree. No special exams for Research Orientation STK796.

b) Masters degree: None of the modules will be repeated more than once. No special exams for Research Orientation STK899.

3) A student who fails one or more modules will be excluded. In this case, the student has the opportunity to lodge an appeal with the relevant faculty.

4) Sick exams will be scheduled for January of the following year. All sick cases must be approved by the Head of the Department of Statistics in writing. No special exams are made available after a sick exam.

Notes

- a. Attendance of the preparation course prior to the commencement of the academic year for the honours research component, if not completed during the honours programme, at the start of the Master's program is compulsory. This is part of the module STK 899 Research Orientation.
- b. Allocation of a supervisor will be in the first semester of first year of study. Kindly make an appointment with the Head of Department. A mini-research proposal may be required for submission in the first year of study prior to continuation of studies in the second year.
- c. It is compulsory for all the postgraduate students to comply with requirements as indicated in STK 899.
- d. Study financing remains the responsibility of the student. The department is not responsible for the funding or obtaining funding for a student. A student must thus plan their finances personally. The department will assist in directing students towards bursary opportunities.
- e. In the case of a dispute with the allocated supervisor(s), the department will proceed as follows:
 - The student must raise their valid concerns in an open professional manner, without fear of retribution.
 - The supervisor(s) has to find constructive solutions to the problem and obtain agreement from the student for the way forward.
 - The above- mentioned process assumes that the student did not commit an act which involves disciplinary actions, such cases should be referred to the legal department.
 - If the conflict/dispute cannot be solved amicably between student and supervisor then either party can request a change in supervisor. If the current promotor agrees **and** a suitable 'new' supervisor can be found then the student may continue with their existing studies. Allocation of a new supervisor is subject to capacity constraints of the department and the current research topic cannot be guaranteed. All existing financial commitments shall be honoured.
 - Finally, in absence of the above-mentioned agreement the student must terminate their current studies.
- f. A student should be available for exams for the entire official exam session in both June and November.
- g. The MEMORANDUM OF UNDERSTANDING (obtainable from clickUP) must be signed by the candidate, the supervisor(s) and the department's postgraduate committee and be submitted to the office of the Head of Student Administration within two months after the date of registration for the research component of the programme. This document must first go through the departments' postgraduate committee. The Department of Statistics requires the following additional procedures for this document:
 - (i) Under supervisor's expectations the following additional requirements should be added:

- The regular scheduled meetings must occur a minimum of every 2 months More frequent meetings are highly recommended to ensure throughput.
- The student must ensure sufficient progress is made so that the supervisor can report positively to the postgraduate committee biannually.
- (ii) Under candidate's expectations the following additional requirements should be added:
 - The supervisor must submit reports to the postgraduate committee on the student's progress biannually to ensure visibility of the student's progress.
- h. On completion of the research component of the Master's degree a student must present a final seminar in the department or at a conference before submission of the dissertation.
- i. Tests and assignments may be taken into account in the determination of the final mark in each of the lectured modules. Exam entrance may be enforced in these modules on discretion of the lecturer.
- j. Examinations in modules take place during one of the official university examination time periods.
- k. Title registration has to be submitted to the Faculty Administration 6 months prior to submission of the mini dissertation.
- I. The research component must be handed in together with an acceptable TurnItIn report.

Programme composition (Total credits required: 180)

A candidate must compile their curriculum in consultation with the department's postgraduate committee. The Master's degree is awarded on the grounds of a dissertation and additional prescribed module work. A pass mark must be obtained in the dissertation as well as in the additional module work.

For:

MSc in Advanced Data Analytics
WST895 Mini-dissertation: Mathematical statistics (100)
STK899 Research orientation (0)

And choose four of the following modules: WST802 Cyber Analytics (20) MVA880 Statistical and machine learning (20) STK880 Capita selecta: Statistics (20) TRA880 Analysis of time series (20) TRG880 Data Science: analytics and visualisation (20)

MCom in Advanced Data Analytics
STK895 Mini-dissertation (100)
STK899 Research orientation (0)

And choose four of the following modules: WST802 Cyber Analytics (20) MVA880 Statistical and machine learning (20) STK880 Capita selecta: Statistics (20) TRA880 Analysis of time series (20) TRG880 Data science: analytics and visualisation (20)

 MSc eScience Curriculum: Year 1 Minimum Credits: 90 Core modules

NEP 801 Research methods and capstone project in data science 801 (15) NEP 802 Data privacy and ethics 802 (15)

Elective modules

Select four of the following modules: NEP 803 Adaptive computation and machine learning 803 (15) NEP 804 Data visualisation and exploration 804 (15) NEP 805 Large-scale computing systems and scientific programming 805 (15) NEP 806 Mathematical foundations of data science 806 (15) NEP 807 Special topics in data science 807 (15) NEP 808 Statistical foundations of data science 808 (15) NEP 809 Large-scale optimisation for data science 809 (15)

Curriculum: Year 2

Minimum Credits: 90 Research NEP 800 Mini-dissertation: eScience 800 (90)

Particulars regarding the contents of modules (MCom and MSc)

Data science: analytics and visualisation (TRG 880)

Prescribed book:

- Hastie, T; Tibshirani, R and Friedman, J: The elements of Statistical learning (Second edition) (Springer, 2009)
- Bishop, C: Machine learning and pattern recognition (Springer, 2006)
- Bouveyron, C; et. al: Model-based clustering and classification for data science (Cambridge, 2019)

Curriculum:

Supervised learning and applications. Multicollinearity, ridge regression, the LASSO and the elastic net. Parametric and nonparametric logistic regression and nonlinear regression Survival regression. Regression extensions: Random forests MARS and Conjoint analysis. Neural networks.

Statistical and machine Learning (MVA 880)

Prescribed book:

- Paolo Giudici : Applied Data Mining (John Wiley ,2003)
- Hastie, T; Tibshirani, R and Friedman, J: The elements of Statistical learning (Second edition) (Springer, 2009)

Curriculum:

Unsupervised learning: deterministic clustering, model-based clustering, latent class and

behavioural analytics, dimension reduction, Topic modelling. Organisation of data, data wrangling and data structure exploration.

- Capita Selecta : Statistics (STK 880)

Curriculum:

This module current and contemporary research topics in advanced data analytics.

- Time Series Analysis (TRA 880)

Curriculum:

Difference equations. Lag operators. Stationary ARMA processes. Maximum likelihood estimation. Spectral analysis. Vector processes. Non-stationary time series. Long-memory processes.

Cyber Analytics (WST 802)

Curriculum:

Reviewing, from a statistical perspective, the cyber-infrastructure ecosystem including distributed computing, multi node and distributed file eco systems, such as Amazon Web Services. Structured and unstructured data sources, including social media data and image data. Setting up of large data structures for analysis. Algorithms and techniques for computing statistics and statistical models on distributed data. Software to be used include, Hadoop, Map reduce, SAS, SAS Data loader for Hadoop.

Research orientation 899 (STK 899)

Curriculum:

The yearbook indicates that to comply with requirements for this module, students need to

- 1. Attend the research orientation bootcamp (which was hosted in February);
- 2. Attend 3 departmental/other approved seminars; and

3. Be registered for the research orientation module each year of registration for the postgraduate degree in question.

Regarding point 2 above:

Students will be expected from 2023 onwards (from now) to attend at least 3 departmental seminars (in person/online where available) and submit a 2-3-page report in October each year of registration via clickUP with a brief summary of each seminar with regards to the topic, speaker, and aspects of the topic presented at the seminar.

Submissions will then be taken into consideration to recommend that a student complies with requirements for the research orientation module.

Note that:

Even if you only complete coursework this year and complete your minidissertation in the following year, you still need to be registered and comply with requirements for the research orientation module in each year of registration.

If you do not comply with requirements for the research orientation module, you will be unable to graduate as the research orientation module forms part of the yearbook description of the degree that you are enrolled for.

If you have already attended and complied with requirements for attendance of the research orientation bootcamp (for example, in your honours year) you need not attend it again in a subsequent year/for a subsequent postgraduate degree in this department.

If you attend and present at a conference/symposium on work regarding your research component, you can include this as part of the summary.

Mini-dissertation: eScience 800 (NEP 800)

Curriculum:

This is the research component of the MSc (eScience) degree and comprises a minidissertation which develops the research skills and bridges the gap between theory and practice.

Prerequisite: Completion of the coursework for the programme

Research methods and capstone project in data science 801 (NEP 801)

Curriculum:

Scientific writing styles; layouts for assignments, projects, theses or publications; research methodologies; scientific assignments; integration of all the aforementioned content items for a capstone project in data science.

Period of presentation Semester 1 or 2

Data privacy and ethics 802 (NEP 802)

Curriculum:

Technical processes of data collection, storage, exchange and access; Ethical aspects of data management; Legal and regulatory frameworks in South Africa and in relevant jurisdictions; Data policies; Data privacy; Data ownership; Legal liabilities of analytical decisions and discrimination; and the Technical and algorithmic approaches to enhance data privacy, and relevant case studies.

Period of presentation Semester 1 or 2

Adaptive computation and machine learning 803 (NEP 803)

Curriculum:

Introduction: Basic concepts. Supervised learning setup: Least means squares, logistic regression, perceptron, exponential family, generative learning algorithms, Gaussian discriminant analysis, naïve Bayes, support vector machines, model selection and feature selection. Learning theory: bias/variance tradeoff, union and Chernoff/Hoeffding bounds, VC dimension, worst case (online) learning. Unsupervised learning: clustering, k-means, expectation maximization, mixture of Gaussians, factor analysis, principal components analysis, independent components analysis. Reinforcement learning and control: Markov decision processes, Bellman equations, value iteration and policy iteration, Q-learning, value function approximation, policy search, reinforce, partially observable Markov decision problems.

Period of presentation Semester 1 or 2

Data visualisation and exploration 804 (NEP 804)

Curriculum:

Data and image models; visualisation attributes (colour) and design (layout); exploratory data analysis; interactive data visualisation; multidimensional data; graphical perception; visualisation software (Python & R); and types of visualisation (animation, networks and text). <u>Period of presentation</u> Semester 1 or 2

Large-scale computing systems and scientific computing 805 (NEP 805)

Curriculum:

Introduction to scientific computing architectures in Python, introduction to distributed systems, introduction to distributed databases, introduction to parallelism, large-data computation and storage models, introduction to well-known distributed systems architectures, and programming large-data applications on open-source infrastructures for data processing and storage systems.

Period of presentation Semester 1 or 2

Mathematical foundations of data science 806 (NEP 806)

Curriculum:

High-dimensional space, best-fit subspaces and singular value decomposition, random walks and Markov chains, statistical machine learning, clustering, random graphs, topic models, matrix factorisation, hidden Markov models, graphical models, wavelets, and sparse representations.

Period of presentation Semester 1 or 2

Special topics in data science 807 (NEP 807)

<u>Curriculum:</u> Specialised and applied concepts and trends in data science. <u>Prerequisites</u> Completion of the coursework of the programme. <u>Period of presentation</u> Semester 1 or 2

Statistical foundations of data science 808 (NEP808)

Curriculum:

An understanding of multivariate statistics, hypothesis testing and confidence intervals. The ability to model data using well-known statistical distributions as well as the ability to handle data that is both continuous and categorical. The ability to perform statistical modelling including multivariate linear regression and adjust for multiple hypotheses. Forecasting, extrapolation, prediction and modelling using statistical methods. Bayesian statistics, an understanding of bootstrapping and Monte Carlo simulation.

Period of presentation Semester 1 or 2

Large-scale optimisation for data science 809 (NEP 809)

Curriculum:

Introduction to convex optimisation, subgradient methods, decomposition and distributed optimisation, proximal and operator splitting methods, conjugate gradients, and nonconvex problems.

Period of presentation Semester 1 or 2

Doctoral Programme

General

The Doctoral degree is presented in the following disciplines: Mathematical Statistics and Statistics. The study is performed under the supervision of a promoter/supervisor. The topic is determined and the promoter is appointed in consultation with and subject to the approval of the head of the department and the department's postgraduate committee. The topic of the thesis to be submitted must be approved by the Dean on recommendation of the head of the department. The Mathematical Statistics option involves research on theoretically identified aspects with applied demonstrations and examples of the concepts. The Statistics option involves research into theoretical aspects identified through specific application needs and requirements. The Statistics option involves more application based research with some theoretical work, that is, new statistical techniques for a specific application.

Admission Requirements

- a. A relevant Master's degree in Mathematical Statistics (Advanced Data Analytics) or Statistics (Advanced Data Analytics) (or reasonable/obvious equivalent, such as Data Science/Biostatistics) is required.
- b. Applicants' obtained Masters degree should include a research dissertation and a full set of course work that consist of core topics that underpin further study within the discipline of statistics.
- c. For PhD (Mathematical Statistics) an average mark of at least 65% in the modules of the MSc Advanced Data Analytics/Mathematical Statistics, with a minimum of 65% for the dissertation.
- d. For PhD (Mathematical Statistics) and PhD (Statistics) an average mark of at least 65% in the modules of the MCom Mathematical Statistics/Statistics/Advanced Data Analytics, with a minimum of 65% for the dissertation.
- e. Before admission to the program and after being shortlisted a prospective student is required to submit (1) a detailed research proposal, (2) present a seminar to the postgraduate committee of the department; and (3) submit a letter of intent. Any of these requirements may be waived at the discretion of the Head of Department. The latter should address the following:

Why the applicant wishes to enroll for a PhD?

What field of study in statistics the applicant has been exposed to (time series / distribution theory etc.)?

Which publications/proceedings (if any) you already have?

What field has the applicant identified, and what reading (general and academic journals) has been done to convince the student to enroll for a PhD?

Possibility of identified supervisor?

Interdisciplinary topics, or work with other academic departments or universities that is envisaged for the degree?

Current work situation: full time student, or part time student and full time employee elsewhere?

Declaration of understanding that the PhD degree weighs a total (minimum) of 360 credits and understanding of how much work this entails (time wise and volume-wise). Understanding the faculty regulations of one paper accepted (NAS)/submitted (EMS) from PhD research before graduation can occur.

Admission to the program will based on the outcome of (1) - (3). A Skype session for the seminar can be organized if required.

- f. Please see the additional requirements for external students as well as international students at the end of this document. Student numbers are limited to a maximum of 10, collectively over all Doctoral Programs in the Department of Statistics.
- g. Historical performance during prior studies will also be considered in selecting students. Specific attention will be given to modules repeated and duration of study.
- h. Admission is also subject to the availability of supervisory capacity in the field of specialisations available in the Department.
- i. Additional entrance requirements as specified by the head of the department. Students with gaps in between studies may be required to write an entrance evaluation.
- j. A student may only commence studies in this department at the beginning of an academic year.

Duration

As long as progress is satisfactory, renewal of registration of a doctoral student will be accepted for a second year and third year of study in the case of a full-time student. Renewal of registration for a fourth and subsequent years for a full-time student will only be considered when Student Administration of the Faculty receives a written motivation (including a time line (Gantt chart) indicating the expected completion date) that is supported by the head of department and department's postgraduate committee.

Promotion

The progress of all doctoral candidates is monitored biannually by the supervisor and the postgraduate committee. A candidate's study may be terminated if the progress is unsatisfactory or if the candidate is unable to finish their studies during the prescribed period. The regulations of the university with regard to doctoral degree studies apply.

Programme composition

A candidate must complete a thesis in one of several fields in Mathematical Statistics or Statistics in which research is actively being done within the Department. The thesis must be handed in together with an acceptable TurnItIn report.

In addition the module STK 911 has to be completed (0 credits).

STK911 description:

The yearbook indicates that to comply with requirements for this module, students need to

1. Attend the research orientation bootcamp (which was hosted in February);

2. Attend 3 departmental/other approved seminars; and

3. Be registered for the research orientation module each year of registration for the postgraduate degree in question.

Regarding point 2 above:

Students will be expected from 2023 onwards (from now) to attend at least 3 departmental seminars (in person/online where available) and submit a 2-3-page report in October each year of registration

via clickUP with a brief summary of each seminar with regards to the topic, speaker, and aspects of the topic presented at the seminar.

Submissions will then be taken into consideration to recommend that a student complies with requirements for the research orientation module.

Note that:

You need to be registered and comply with requirements for the research orientation module in each year of registration.

If you do not comply with requirements for the research orientation module, you will be unable to graduate as the research orientation module forms part of the yearbook description of the degree that you are enrolled for.

If you have already attended and complied with requirements for attendance of the research orientation bootcamp (for example, in your honours year) you need not attend it again in a subsequent year/for a subsequent postgraduate degree in this department.

If you attend and present at a conference/symposium on work regarding your research component, you can include this as part of the summary.

Notes

Attendance of the preparation course prior to commencement of the academic year for the honours research component, if not completed during the honours or Masters programme, at the start of the Doctoral program is compulsory. This forms part of STK 911.

Study financing remains the responsibility of the student. The department is not responsible for the funding or obtaining funding for a student. A student must thus plan their finances personally. The department will assist in directing students towards bursary opportunities.

It is compulsory for all the postgraduate students to comply with requirements as indicated in STK 911.

In the case of a dispute with the allocated supervisor(s), the department will proceed as follows:

- The student must raise their valid concerns in an open professional manner, without fear of retribution.
- The supervisor(s) has to find constructive solutions to the problem and obtain agreement from the student for the way forward.
- The above- mentioned process assumes that the student did not commit an act which involves disciplinary actions, such cases should be referred to the legal department.
- If the conflict/dispute cannot be solved amicably between student and supervisor then either party can request a change in supervisor. If the current promotor agrees **and** a suitable 'new' supervisor can be found then the student may continue with their existing studies. Allocation of a new supervisor is subject to capacity constraints of the department and the current research topic cannot be guaranteed. All existing financial commitments shall be honoured.
- Finally, in absence of the above-mentioned agreement the student must terminate their current studies.
- A student should be available for exams for the entire official exam session in both June and

November.

The MEMORANDUM OF UNDERSTANDING (available from clickUP) must be signed by the candidate, the supervisor(s) and the postgraduate committee and be submitted to the office of the Head of Student Administration within two months after the date of registration for the research component of the programme. This document must first go through the departments' Postgraduate Committee. The Department of Statistics requires the following additional procedures for this document:

(i) Under supervisor's expectations the following additional requirements should be added:

- a. The regular scheduled meetings must occur a minimum of every 2 months (not the 3 months indicated in the document). More frequent meetings are highly recommended to ensure throughput.
- b. The candidate is required to present at least 2 seminars a year to the department and its students on the progress of the work.
- (ii) The student must ensure sufficient progress is made so that the supervisor can report positively to the postgraduate committee biannually.
- (iii) Under candidate's expectations the following additional requirements should be added:
 - a. The supervisor must submit reports to the postgraduate committee on the student's progress biannually to ensure visibility of the student's progress.

Masters and Doctoral degree research studies focus areas:

Only research topics within the department's <u>research fields</u> can be accommodated. View the departmental website/staff profiles for more information.

Additional requirements for students entering UP from externally (including international students)

- 1. A compulsory language proficiency test must be completed at UP. The postgraduate committee of the department will help with the facilitation of the test through the universities language unit. It may be required, based on the outcome, that a student does additional language courses.
- 2. Applications to study at UP must include full academic transcripts from undergraduate to current level, include SAQA accreditation (for international students only but may be required for South African students on request). The previous degree should have compatible content with the degree applied for at the University of Pretoria.
- 3. An entrance evaluation will be required of all external applicants.
- 4. A research proposal and/or previous research work must be submitted with the application, together with a self-evaluation (1/2 page).

Further information

There are excellent opportunities for postgraduate bursaries. For postgraduate bursaries or any other information contact the Department of Statistics. See also departmental website. Bursaries are on based on academic merit, academic performance, academic progress in the programme and availability of funds, however funding support remains the responsibility of the student. The Main Campus library provides excellent postgraduate study facilities.

All postgraduate programme inquiries please contact Prof JT Ferreira: Johan.Ferreira@up.ac.za

Contact details:

Department of Statistics, University of Pretoria, PRETORIA, 0002 +27 12 420-3774 or +27 12 420-3450 Website: Department of Statistics at <u>http://www.up.ac.za/statistics</u>

Young alumni have their say

"I build statistical models to help banks make sound decisions about who they grant loans to and whether a customer can actually afford the credit they are applying for. Statistics is a thrilling field with so many applications in every industry. Statistics helps to answer questions. Statistics is a rare skill that is highly sought after in the workplace."

Buwang Mokuele (Decision Analyst)

"People across the globe need statistical skills. Statistics is everywhere and relates to everyday life in countless ways. News, information exchanges, interactions and events involve statistics. I am grateful to be a statistician!"

Seite Makgai (Lecturer in the Department of Statistics at UP)