Statistics for the College of Science, Engineering and Technology

(offered by the Department of Statistics)
Telephone number 012 429 6464

1 Introduction

The primary objective of Statistics is to analyse data. Analyses of this type usually involve the formulation of mathematical models, and therefore the more mathematical aspects of statistics should not be neglected. Statistical modelling and analysis today play an important role in many areas of research (medicine, pharmacy, agriculture, psychology, sociology, engineering, the chemical industry and many other areas) as well as in the processing of management information for commerce and industry.

In the College of Science, Statistics may be taken as a major or it can be combined with other subjects. Streams within the BSc specialisation degrees in which Statistics is a major are:
- Mathematical Statistics
- Statistical Decision Science
- Statistics for Management
- Environmental Statistics
- Computational Statistics (Programming)
- Computational Statistics (Information Systems)

Full particulars of the various streams appear in Part 7 of the Calendar (College Calendar). Statistics can also be taken as a subject in the College of Economic and Management Sciences and in the College of Human Sciences (see relevant College Calendar).

If you prefer not to choose any of the streams within the BSc specialisations, you may compile your own curriculum. This amounts to making a particular choice of modules in accordance with section 3 (Statistics as a major subject) given below.

1.1 Statistics and other BSc subjects

Statistics is used extensively in all scientific fields and therefore the study of Statistics may usefully be combined with any BSc subject. If you are interested in a career in science, you should consider combining Statistics with Mathematics, Computer Science or Operations Research.

We recommend that, before you compile your own curriculum, you should consider the streams within the specialisation degrees which are described below: (the curricula of these streams appear in the College Calendar).

(a) Mathematical Statistics Stream
Statistics is the collection and analysis of data, followed by the interpretation and presentation of the information in the data. This stream develops a thorough mathematical foundation on which statistical theories are built. It also prepares the graduate for postgraduate studies in Statistics.

(b) Statistical Decision Science Stream
This stream provides the graduate with knowledge of operations research, statistics and aspects of computer science and informatics. It also prepares the graduate for postgraduate studies in data mining, operations research and statistics.

(c) Statistics for Management Stream
This stream develops programming, modelling, simulation, risk analysis, forecasting and statistical techniques required to solve problems in manufacturing, banking, business management and management consultancy. It also prepares the graduate for postgraduate studies in Statistics and/or Operations Research.

(d) Environmental Statistics Stream
This stream provides the statistical and geographical knowledge required to solve problems in environmental evaluation, population, spatial structures and the interpretation of aerial photos. It also prepares the graduate for postgraduate studies in Statistics.

(e) Computational Statistics (Programming) Stream
Statistics requires the intelligent, critical and judicious use of computers. This stream provides the necessary statistical and computational techniques and skills required to make a contribution to a computing team. It also prepares the graduate for postgraduate studies in Statistics.

(f) Computational Statistics (Information Systems) Stream
This stream aims to educate and train statistical practitioners and professionals primarily involved in the business information needs of an organisation. It provides a sound understanding of computer databases, and the essence of statistical theory. It also prepares the graduate for postgraduate studies in Statistics.
1.2 Requirements for admission to postgraduate studies

To qualify for admission to studies for the Honours BSc degree in Statistics a student must:

- hold a Bachelor’s degree with a major in Statistics (which implies a significant mathematics component)
- final year Statistics modules/courses must have been passed with an average of at least 60% within the past 5 years
- have passed MAT211 and MAT215 or the equivalent
- have access to a computer as CDs form part of the study material in certain modules.

Full particulars regarding prerequisites, syllabuses and the compilation of a curriculum are contained in a separate brochure on postgraduate studies in Statistics, which is obtainable free of charge from the Registrar on application.

NB
Admission to studies for the Honours BSc degree rests with the Department.

2 General Information

- From 2007 all modules are linked to a year mark system. The year mark plus the October/November examination mark will be your final mark. Should you have a supplementary examination, the year mark is not incorporated into that examination mark. In Statistics the year mark is calculated according to marks earned in assignments. More details will be given, per module, in Tutorial letter 101.
- A knowledge of Mathematics is absolutely essential for the statistician. The use of a non-programmable pocket calculator is permissible in the examination.
- Access to a computer is compulsory from the second level onwards as CDs form part of the study material in certain modules.
- Credit for a BSc degree is granted for:
  (i) either STA102 (prior to 2001) or STA105
  (ii) either STA112 or STA106
  (iii) either STS101 or STA105
  (iv) either STS101 and STS102 or STA101, 105, 106 and 111
  (v) either STS111 and STS112 or STA121, 122, 123 and 124
  (vi) either STS111 or STA121 and 122
  (vii) either STS112 or STA123 and STA124
  (viii) either STA101 or STA222
  (ix) either STA105 or STA121
  (x) either STA106 or STA124
  (xi) credits for other previously passed Statistics courses is at the discretion of the Department.

3 Transitional arrangements

- Students who passed STA101, will receive credit for STA122.
- Students who passed STA105, will receive credit for STA121.
- Students who passed STA106, will receive credit for STA124.

4 Statistics as a Major Subject

**Compulsory modules for a major subject combination:**

Prerequisite to enroll for a major in statistics, ONE of the following:

(a) Mathematics HIGHER GRADE passed with at least 50% (D symbol) or 80% (A symbol) on STANDARD GRADE at Matriculation level.
(b) Mathematics at Matriculation level passed prior to level differentiation with a corresponding pass percentage.
(c) An equivalent examination in Mathematics with a corresponding pass percentage.

If you do not satisfy this higher mathematics requirement, you may enroll for the module MAT110 [with prerequisite Sc1(1)(b)(i)-(iv) in Calendar 7] and present successful completion of it for admission into a BSc with a major in Statistics.

*First level:* STA121, 122, 123 and STA124
  - plus
  - MAT103, 111, 112 and MAT113

*Second level:* STA202, 203 and at least TWO of STA204, 206, 208
  - plus
  - MAT211 and 215

*Third level:* STA302, 303, 305 and at least ONE of STA306 or 312

The choice of optional modules depends on the areas of application in which the student intends to work.

**RECOMMENDATIONS:**

- For biological and medical applications, include STA204, 206 and 312.
- For applications in the earth sciences, include STA204.
- For applications in the behavioural sciences, include STA206, 208 and 312.
- For engineering applications, include STA204 and 306.
- For business applications, include STA206, 208 and 312.
Any mathematical, behavioural or experimental science will be a sensible subject to combine with Statistics, depending on the area of application. We mention in particular Mathematics, Applied Mathematics, Computer Science, Operations Research, Astronomy, Chemistry, Physics, Psychology, Geography or any of the Biological Sciences.

NB

Only in the case of certain BSc specialisation degrees are students permitted to take STS111 and STS112 for first-level modules.

5

Syllabus

NB

All modules in this subject are offered as YEAR MODULES.

FIRST-LEVEL MODULES

Advice: We recommend that you study your first-level modules in the order STA121, 122, 123 and 124.

Prerequisite: Mathematics as in Sc1(1)(b) or (c) in Part 7 of the Calendar

STA121M Descriptive statistics (2 hours)*

Purpose: to present and critically analyse visual and numerical representations of data using descriptive statistics. Evaluation of the components and characteristics of different time series as well as fundamental knowledge of indices enables the qualified student to interpret and evaluate economic indicators presented in the media and economic environments. This module will be useful to the future professional statistician in the academic, government or industry sectors to formulate educated decisions based on fundamental knowledge.

STA122N Probability and probability distributions (2 hours)*

Co-requisite: STA121

Purpose: to interpret and apply basic probability concepts and be able to differentiate between probability distributions. The reliability of conclusions about the population based on sample information is evaluated. This module will be useful to the future professional statistician in the academic, government and industry sectors to formulate educated decisions based on fundamental knowledge.

STA123P Statistical inference (2 hours)*

Co-requisite: STA121 and STA122

Purpose: to construct meaning from decision-making and prediction (via sampling and sampling distributions) and their relevance to sample statistics in repeated sampling. Evaluation of the reliability of estimates, conducting statistical tests of hypotheses and constructing confidence intervals, all components in the fundamental knowledge of the future professional statistician in the academic government and industry sectors, form part of the tools for inferential statistics.

STA124G Data analysis (2 hours)*

Co-requisite: STA121, 122 and 123

Purpose: to discover the very effective and widely used technique of analysis of variance (ANOVA) to detect possible differences between the treatment means of two or more populations of interval data. The analysis of categorical data and applicable tests are also explored. Credited students are aware that for ordinal data, the mean is not an appropriate measure of location and characteristics of the population are measured by nonparametric techniques. Students will be introduced to regression analyses for relating variables and its application in forecasting and managerial decision-making and that the strength of the relation is measured by correlation analysis.

STS1113 Descriptive statistics and probability (3 hours)*

Prerequisite: Rule G13, Part 1 of the Calendar

Purpose: to have an informed understanding of exploratory data analysis as used in graphical and tabular techniques; measures of central location, variability and linear relationships; simple sampling procedures. Students will be able to use probability as a tool to create discrete and continuous probability distributions, used extensively in statistical inference. The contents of this module have important applications in finance and are useful in several management sciences.

STS1124 Data analysis and inference (3 hours)*

Prerequisite: Rule G13, Part 1 of the Calendar

Co-requisite: STS111

Purpose: to have a basic perspective of the role of the sampling distribution of the mean, a proportion and the difference between two means in statistical inference, interval estimation and hypothesis testing. Students will be able to estimate single and combinations of population parameters; understand one way analysis of variance; apply parametric and nonparametric tests such as two Chi-squared tests and the Wilcoxon signed rank sum test. They will also be familiar with simple linear regression and correlation, as well as with the basics of time series analysis and forecasting. The contents of this module are relevant in a wide variety of applications in business and economics and represent a significant contribution to the development of the student as a statistics practitioner.

STS1055 Basic statistics (3 hours)*

Prerequisite: Rule G13, Part 1 of the Calendar

Purpose: to gain an understanding of basic statistical concepts and turn data into information. After completion students will be able to distinguish between the different types of data and their summaries, will have knowledge of different sampling methods and surveys; can examine relationships between quantitative variables. They will be able to interpret probability concepts and apply the rules; differentiate between discrete and continuous random variables with the binomial and normal random variables as respective examples; understand sampling distributions for one mean and one proportion as well as for the difference of two means and two proportions – all as preparation for appreciation of statistical inference. Students will acquaint themselves
SECOND-LEVEL MODULES

NB

The module STA211 has been phased out.

STA202M  Distribution theory (3 hours)*
Prerequisite: STA121, 122 or STS111, 112, MAT103, 111
Co-requisite: MAT211
Purpose: to gain insight into the role that formal theory plays in data analytic methods, discussing a wide variety of discrete and continuous distributions simultaneously. After completion students should understand the joint probability structure of two random variables (discrete and continuous case); be able to calculate expectation, variance, covariance, conditional expectation and moment-generating functions; have insight into distributions of functions of independent random variables; prove the law of large numbers and the central limit theorem under fairly strong assumptions; comprehend how the Chi-square, t, and F distributions are derived from the normal distribution.

STA203N  Applied statistics (3 hours)*
Prerequisite: STA101 or 105 or STA121, 122, 123 or STS111, 112. Access to a computer is compulsory.
Purpose: to enable students to identify the correct technique, manage the statistical software JMP to do the computations and interpret the results for decisions regarding tests for normality, independence and hypotheses concerning means, variances and regression.

STA204P  Design of experiments (3 hours)*
Prerequisite: STA105 or 106 or STA123, 124 or STS111, 112
Purpose: to enable students to gain a basic understanding of the design of experiments through an introduction to analysis of variance; elementary concepts; factorial designs; fractional designs, and more complicated designs.

STA206R  Sampling and survey methods (3 hours)*
Prerequisite: STA105 or STA123 or STS111
Purpose: to gain insight into the practical aspects of survey problems, focusing on applications; sample survey design; estimation procedures and the fundamental role that probability plays in making inferences. After completion students should have appreciation for the importance of questionnaire design, methods of data collection, estimation procedures and sources of errors in surveys.

THIRD-LEVEL MODULES

STA302Q  Distribution theory (3 hours)
Prerequisite: STA202 or (STA203 and QMS101)
Co-requisite: MAT211 or STA311
Purpose: to gain insight into distributions and their relationships. After completion students should comprehend noncentrality; understand compounding and generalisation as methods for finding parameter-rich distributions; use bivariate and multivariate distributions to describe normal and non-normal variables.

STA303R  Inference (3 hours)
Prerequisite: STA202
Co-requisite: MAT215 or STA311
Purpose: to gain insight into likelihood, data reduction, point estimation, and interval estimation.

STA305T  Analysis of variance and regression (3 hours)
Prerequisite: STA203 plus any other second-level Statistics module. Access to a computer is compulsory.
Co-requisite: STA311 and STA313
Purpose: to enable students to demonstrate an understanding of one- and two-way analysis of variance, fixed effects and mixed models, and simple and multiple linear regression.

STA306U  Sampling techniques (3 hours)*
Prerequisite: STA203 plus any other second-level Statistics module.
Advice: Knowledge of the content of STA206 is advised.
Purpose: to gain more advanced insight into stratified random sampling; systematic and cluster sampling; estimation of the sample size; ratio and regression estimation; sampling with unequal probabilities; complex surveys; non-response.

STA312S  Time series (3 hours)*
Prerequisite: STA203 plus any other second-level Statistics modules
Corequisite: STA303
Advice: This module is a continuation of STA208.
Purpose: to gain insight into Box-Jenkins methodology; AR, MA and ARIMA models; also to use statistical software for practical modelling of time series.