REGULATIONS FOR THE DEGREE OF
MASTER OF STATISTICS
(MStat)
For students admitted in 2019-2020 and thereafter

(See also General Regulations and Regulations for Taught Postgraduate Curricula)

Any publication based on work approved for a higher degree should contain a reference to the effect that the work was submitted to the University of Hong Kong for the award of the degree.

Admission requirements

MS 1.   To be eligible for admission to the courses leading to the degree of Master of Statistics a candidate
  (a) shall comply with the General Regulations and the Regulations for Taught Postgraduate Curricula;
  (b) shall hold
    (i) a Bachelor's degree with honours of this University, or
    (ii) another qualification of equivalent standard from this University or another University or comparable institution acceptable for this purpose; and
  (c) shall pass a qualifying examination if so required.

Period of study

MS 2.   The curriculum shall normally extend over one academic year of full-time study or two academic years of part-time study. Candidates shall not be permitted to extend their studies beyond the maximum period of registration of two academic years of full-time study or three academic years of part-time study, unless otherwise permitted or required by the Board of the Faculty.

Exemption

MS 3.   A candidate with appropriate qualification and professional experiences may, on production of appropriate certification of having satisfactorily completed another course or courses equivalent in content to any of the two compulsory courses as specified in the syllabuses, be exempted from the compulsory course(s), subject to approval of the Board of the Faculty.   Candidates so exempted must replace the number of exempted credits with electives course(s) in the curriculum of the same credit value.

Award of degree

MS 4.   To be eligible for the award of the degree of Master of Statistics, a candidate shall
  (a) comply with the General Regulations and the Regulations for Taught Postgraduate Curricula;
  (b) successfully complete the curriculum in accordance with the regulations set out below; and
  (c) have achieved a cumulative grade point average of at least 2.0.

Any student who does not meet the exit requirement 4(c) has neither fulfilled the requirements for graduation or discontinuation. Therefore, the students concerned would be allowed to continue their study until the end of the maximum study period as outlined in the regulations.

Completion of curriculum

MS 5.   To successfully complete the curriculum, a candidate shall satisfy the requirements prescribed in TPG 6 of the Regulations for Taught Postgraduate Curricula; follow courses of instruction; and satisfy the examiners in the prescribed courses and in any prescribed form of examination in accordance with the regulations set out below.

Assessments

MS 6.   (a)   In any course where so prescribed in the syllabus, coursework or a project report may constitute part or whole of the examination for the course.
         (b)   Except where otherwise stated, the assessment weight assigned to coursework will be 25% of the final grade of each course.

MS 7.   If during any academic year a candidate has failed at his/her first attempt in a course or courses, but is not required to discontinue his/her studies by Regulation MS 8, the candidate may be permitted to make up for the failed courses in the following manner:
         (a)   undergoing re-assessment/re-examination in the failed course or courses to be held before the next academic year, resulting in no more than a D grade or a pass grade under this provision; or
         (b)   for repeating the course and re-examination in the failed course or courses in the next academic year; or
         (c)   for elective courses, taking another course in lieu and satisfying the assessment requirements.

MS 8.   Failure to undertake the examination of a course as scheduled shall normally result in automatic failure in that course. A candidate who, because of illness, is unable to be present at the written examination of any course may apply for permission to present himself/herself at a supplementary examination of the same course to be held before the beginning of the following academic year. Any such application shall be made on the form prescribed within two weeks of the first day of the candidate's absence from any examination. A supplementary examination authorized under these circumstances shall not be subject to Regulation MS 7(a).

MS 9.   A candidate who
         (a)   during any academic year has failed in half or more than half the number of credits of all the courses to be examined in that academic year; or
         (b)   has failed at a repeated attempt in any course; or
         (c)   has exceeded the maximum period of registration

may be required to discontinue his/her studies.

Grading system

MS 10.   Individual courses shall be graded according to one of the following grading systems as determined by the Board of Examiners:

(a) Letter grades, their standards and the grade points for assessment as follows:

Grade	Standard	Grade Point
A+	**Excellent**	4.3
A		4.0
A-		3.7
B+	**Good**	3.3
B		3.0
B-		2.7
C+	**Satisfactory**	2.3
C		2.0
C-		1.7
D+	**Pass**	1.3
D		1.0
F	**Fail**	0

or

(b) 'Distinction', 'Pass' or 'Fail'.

Courses which are graded according to (b) above will not be included in the calculation of the GPA.

Assessment results

MS 11.   On successful completion of the curriculum, candidates who have shown exceptional merit at the whole examination may be awarded a mark of distinction, and this mark shall be recorded in the candidates' degree diploma.

**SYLLABUSES FOR THE DEGREE OF
MASTER OF STATISTICS**

The Department of Statistics and Actuarial Science offers a postgraduate curriculum leading to the degree of Master of Statistics, with two study modes: the one year full-time mode and the two years part-time mode. The curriculum is designed to provide graduates with training in the principles and practice of statistics. Candidates should have knowledge of matrices and calculus. Candidates applying for full-time mode of study should also have knowledge of introductory statistics and linear modelling.

The curriculum offers great flexibilities for students who wish to take a general approach or a specialized theme in Risk Management or Data Analytics. A student may choose to have his/her theme printed on the transcript if he/she has satisfied one of the theme requirements.

STRUCTURE AND EVALUATION

Each student must complete at least 60 credits of courses. Courses with 6 credits are offered in the first and second semesters while courses with 3 credits may be offered in the winter break and summer semester. If a student selects an MStat course whose contents are similar to a course (or courses) which he/she has taken in his/her previous study, the Department may not approve the selection in question.

Curriculum for full-time study

The period of study extends over one year full-time commencing September. Among the 60 credits of courses, students should take two compulsory courses: STAT6008 Advanced statistical inference and STAT6014 Advanced statistical modelling.

Two compulsory courses (12 credits)

STAT6008   Advanced statistical inference (6 credits)
STAT6014   Advanced statistical modelling (6 credits)

Theme-specific elective courses (24 Credits)

Risk Management theme	*Data Analytics theme*
plus 24 credits from	**plus 24** credits from
STAT6013 Financial data analysis (6 credits)	STAT6011 Computational statistics (6 credits)
STAT6015 Advanced quantitative risk management and finance (6 credits)	STAT6016 Spatial data analysis (6 credits)
STAT6017 Operational risk and insurance analytics (6 credits)	STAT7005 Multivariate methods (6 credits)
STAT8003 Time series forecasting (6 credits)	STAT7007 Categorical data analysis (6 credits)
STAT8007 Statistical methods in economics and finance (6 credits)	STAT7008 Programming for data science (6 credits)
STAT8014 Risk management and Basel Accords (6 credits)	STAT8003 Time series forecasting (6 credits)
STAT8015 Actuarial statistics (6 credits)	STAT8016 Biostatistics (6 credits)
STAT8017 Data mining techniques (6 credits)	STAT8017 Data mining techniques (6 credits)
STAT8020 Quantitative strategies and algorithmic trading (6 credits)	STAT8019 Marketing analytics (6 credits)
STAT8021 Big data analytics (6 credits)	STAT8021 Big data analytics (6 credits)
STAT8305 Bayesian statistics (3 credits)	STAT8302 Structural equation modelling (3 credits)
	STAT8305 Bayesian statistics (3 credits)
	STAT8306 Statistical methods for network data (3 credits)
	STAT8307 Text analytics (3 credits)

Other elective courses (18 credits)

plus at least **18** credits from
STAT6009   Research methods in statistics (6 credits)
STAT6010   Advanced probability (6 credits)
STAT7006   Design and analysis of sample surveys (6 credits)
STAT7301   Socio-economic statistics for business and public policies (3 credits)
STAT8000   Workshop on spreadsheet modelling and database management (3 credits)
STAT8300   Career development and communication workshop (3 credits)
STAT8304   Current topics in Statistics (3 credits)
Any theme-specific elective courses
Any capstone courses

Capstone requirement (6 credits)

plus 6 credits from
STAT8002   Project (6 credits)
STAT8017   Data mining techniques (6 credits)
STAT8088   Practicum (6 credits)
STAT8089   Capstone Project (6 credits)

Apart from the two compulsory courses and capstone requirement, candidates may choose not to follow any theme and may take 42 credits of elective courses in any order, whenever feasible.

Curriculum for part-time study

The period of study extends over two years part-time commencing September. Among the 60 credits of courses, students should take the two compulsory courses STAT7003 Foundations of statistics and STAT7004 Linear modelling.

Two compulsory courses (12 credits)

STAT7003   Foundations of statistics (6 credits)
STAT7004   Linear modelling (6 credits)

Theme-specific elective courses (24 Credits)

Risk Management theme	*Data Analytics theme*
plus 24 credits from	**plus 24** credits from
STAT6013 Financial data analysis (6 credits)	STAT6011 Computational statistics (6 credits)
STAT6015 Advanced quantitative risk management and finance (6 credits)	STAT6014 Advanced statistical modelling (6 credits)
STAT6017 Operational Risk and Insurance Analytics (6 credits)	STAT6016 Spatial data analysis (6 credits)
STAT8003 Time series forecasting (6 credits)	STAT7005 Multivariate methods (6 credits)
STAT8007 Statistical methods in economics and finance (6 credits)	STAT7007 Categorical data analysis (6 credits)
STAT8014 Risk management and Basel Accords (6 credits)	STAT7008 Programming for data science (6 credits)
STAT8015 Actuarial statistics (6 credits)	STAT8003 Time series forecasting (6 credits)
STAT8017 Data mining techniques (6 credits)	STAT8016 Biostatistics (6 credits)
STAT8020 Quantitative strategies and algorithmic trading (6 credits)	STAT8017 Data mining techniques (6 credits)
STAT8021 Big data analytics (6 credits)	STAT8019 Marketing analytics (6 credits)
STAT8305 Bayesian statistics (3 credits)	STAT8021 Big data analytics (6 credits)
	STAT8302 Structural equation modelling (3 credits)
	STAT8305 Bayesian statistics (3 credits)
	STAT8306 Statistical methods for network data (3 credits)
	STAT8307 Text analytics (3 credits)

Other elective courses (18 credits)

plus at least **18** credits from

STAT6008   Advanced statistical inference (6 credits)
STAT6009   Research methods in statistics (6 credits)
STAT6010   Advanced probability (6 credits)
STAT7006   Design and analysis of sample surveys (6 credits)
STAT7301   Socio-economic statistics for business and public policies (3 credits)
STAT8000   Workshop on spreadsheet modelling and database management (3 credits)
STAT8300   Career development and communication workshop (3 credits)
STAT8304   Current topics in Statistics (3 credits)
Any theme-specific elective courses
Any capstone courses

Capstone requirement (6 credits)

plus 6 credits from

STAT8002   Project (6 credits)
STAT8017   Data mining techniques (6 credits)
STAT8088   Practicum (6 credits)
STAT8089   Capstone project (6 credits)

Apart from the two compulsory courses and capstone requirement, candidates may choose not to follow any theme and may take 42 credits of elective courses in any order, whenever feasible.

Compulsory Course Replacement

Students with prior background may replace each compulsory course with a more advanced course as shown below:	
Replace ...	**With**
STAT6008 Advanced statistical inference	STAT6009 Research methods in statistics
STAT6014 Advanced statistical modelling	Any other course
STAT7003 Foundations of statistics	STAT7005 Multivariate methods
STAT7004 Linear modelling	STAT6014 Advanced statistical modelling

COURSE DESCRIPTION

STAT6008   Advanced statistical inference (6 credits)

This course covers the advanced theory of point estimation, interval estimation and hypothesis testing. Using a mathematically-oriented approach, the course provides a formal treatment of inferential problems, statistical methodologies and their underlying theory. It is suitable in particular for students intending to further their studies or to develop a career in statistical research. Contents include: (1) Decision problem – frequentist approach: loss function; risk; decision rule; admissibility; minimaxity; unbiasedness; Bayes' rule; (2) Decision problem – Bayesian approach: prior and posterior distributions, Bayesian inference; (3) Estimation theory: exponential families; likelihood; sufficiency; minimal sufficiency; completeness; UMVU estimators; information inequality; large-sample theory of maximum likelihood estimation; (4) Hypothesis testing: uniformly most powerful (UMP) test; monotone likelihood ratio; UMP unbiased test; conditional test; large-sample theory of likelihood ratio; confidence set; (5) Nonparametric inference; bootstrap methods.
Assessment: One 2-hour written examination; 40% coursework and 60% examination

---

STAT6009   Research methods in statistics (6 credits)

This course introduces some statistical concepts and methods which potential graduate students will find useful in preparing for work on a research degree in statistics. Focus is on applications of state-of-the-art statistical techniques and their underlying theory. Contents may be selected from: (1) Basic asymptotic methods: modes of convergence; stochastic orders; laws of large numbers; central limit theorems; delta method; (2) Parametric and nonparametric likelihood methods: high-order approximations; profile likelihood and its variants; signed likelihood ratio statistics; empirical likelihood; (3) Nonparametric statistical inference: sign and rank tests; Kolmogorov-Smirnov test; nonparametric regression; density estimation; kernel methods; (4) Computationally-intensive methods: cross-validation; bootstrap; permutation methods; (5) Robust methods: measures of robustness; M-estimator; L-estimator; R-estimator; estimating functions; (6) Other topics as determined by the instructor.
Assessment: One 2-hour written examination; 25% coursework and 75% examination

---

STAT6010   Advanced probability (6 credits)

This course provides an introduction to measure theory and probability. The course will focus on some basic concepts in theoretical probability which are important for students to do research in actuarial science, probability and statistics. Contents include: sigma-algebra, measurable space, measure and probability, measure space and probability space, measurable functions, random variables, integration theory, characteristic functions, convergence of random variables, Hilbert spaces, conditional expectations, martingales.
Assessment: One 2-hour written examination; 25% coursework and 75% examination

STAT6011   Computational statistics (6 credits)

This course aims to give postgraduate students in statistics a background in modern computationally intensive methods in statistics. It emphasizes the role of computation as a fundamental tool of discovery in data analysis, of statistical inference, and for development of statistical theory and methods. Contents include: Bayesian statistics, Markov chain Monte Carlo methods including Gibbs sampler, the Metropolis-Hastings algorithm, and data augmentation; Generation of random variables including the inversion methods, rejection sampling, the sampling/importance resampling method; Optimization techniques including Newton's method, expectation-maximization (EM) algorithm and its variants, and minorization-maximization (MM) algorithms; Integration including Laplace approximations, Gaussian quadrature, the importance sampling method, Numerical optimization and integration, EM algorithm and its variants, Simulation and Monte Carlo integration, Importance sampling and variance reduction techniques; and other topics such as Hidden Markov models, neural networks, and Bootstrap methods.
Pre-requisites: Students should not be taking or have taken STAT8305 Bayesian statistics or equivalent
Assessment: One 2-hour written examination; 25% coursework and 75% examination

STAT6013   Financial data analysis (6 credits)

This course aims at introducing statistical methodologies in analyzing financial data. Financial applications and statistical methodologies are intertwined in all lectures. Contents include: recent advances in modern portfolio theory, copula, market microstructure and high frequency data analysis.
Assessment: One 2-hour written examination; 40% coursework and 60% examination

STAT6014   Advanced statistical modelling (6 credits)

This course introduces modern methods for constructing and evaluating statistical models and their implementation using popular computing software, such as R or Python. It will cover both the underlying principles of each modelling approach and the model estimation procedures. Topics from: (i) Generalized linear models; (ii) Mixed models; (iii) Kernel and local polynomial regression; selection of smoothing parameters; (iv) Generalized additive models; (v) Hidden Markov models and Bayesian networks.
Assessment: One 2-hour written examination; 50% coursework and 50% examination

STAT6015   Advanced quantitative risk management and finance (6 credits)

This course covers statistical methods and models of importance to risk management and finance and links finance theory to market practice via statistical modelling and decision making. Emphases will be put on empirical analyses to address the discrepancy between finance theory and market data. Contents include: Elementary Stochastic Calculus; Basic Monte Carlo and Quasi-Monte Carlo Methods; Variance Reduction Techniques; Simulating the value of options and the value-at-risk for risk management; Review of univariate volatility models; multivariate volatility models; Value-at-risk and expected shortfall; estimation, back-testing and stress testing; Extreme value theory for risk management.
Assessment: One 2-hour written examination; 25% coursework and 75% examination

STAT6016   Spatial data analysis (6 credits)

This course covers statistical concepts and tools involved in modelling data which are correlated in space. Applications can be found in many fields including epidemiology and public health, environmental sciences and ecology, economics and others.   Covered topics include: (1) *Outline* of three types of spatial data: point-level (geostatistical), areal (lattice), and spatial point process. (2) *Model-based geostatistics*: covariance functions and the variogram; spatial trends and directional effects; intrinsic models; estimation by curve fitting or by maximum likelihood; spatial prediction by least squares, by simple and ordinary kriging, by trans-Gaussian kriging. (3) *Areal data models*: introduction to Markov random fields; conditional, intrinsic, and simultaneous autoregressive (CAR, IAR, and SAR) models. (4) *Hierarchical modelling* for univariate spatial response data, including Bayesian kriging and lattice modelling. (5) *Introduction* to simple spatial point processes and spatio-temporal models. Real data analysis examples will be provided with dedicated R packages such as geoR.
Assessment: One 2-hour written examination; 50% coursework and 50% examination

STAT6017   Operational risk and insurance analytics (6 credits)

This course aims to provide the foundation of operational risk management and insurance. Special emphasis will be put on the analytical and modeling techniques for operational risk and insurance. Contents include fundamentals of operational risk and Basel regulation, loss distribution, estimation of risk models, copula and modeling dependence, insurance and risk transfer for operational risk.
Assessment: One 2-hour written examination; 25% coursework and 75% examination

STAT7003   Foundations of statistics (6 credits)

Motivated by real problems involving uncertainty and variability, this course introduces the basic concepts and principles of statistical inference and decision-making. Ideas developed will include probability modelling, statistical distributions; parametric classes; the likelihood principle; maximum likelihood estimation; likelihood ratio tests; hypotheses testing. (Only under exceptional academic circumstances can this compulsory course be replaced by an elective course.)
Assessment: One 3-hour written examination; 25% coursework and 75% examination

STAT7004   Linear modelling (6 credits)

Much of the analysis of variability is concerned with locating the sources of the variability, and many current statistical techniques investigate these sources through the use of 'linear' models. This course presents a unified theory of such statistical problems including regression; variance and covariance analyses; design of experiments; and their practical implementation with statistical packages. (Only under exceptional academic circumstances can this compulsory course be replaced by an elective course.)
Assessment: One 3-hour written examination; 25% coursework and 75% examination

STAT7005   Multivariate methods (6 credits)

In many disciplines the basic data on an experimental unit consist of a vector of possibly correlated measurements.   Examples include the chemical composition of a rock; the results of clinical observations and tests on a patient; the household expenditures on different commodities. Through the challenge of problems in a number of fields of application, this course considers appropriate statistical models for explaining the patterns of variability of such multivariate data. Topics include: multiple, partial and canonical correlation; multivariate regression; tests on means for one-sample and two-sample problems; profile analysis; test for covariances structure; multivariate ANOVA; principal components analysis; factor analysis; discriminant analysis and classification.
Assessment: One 3-hour written examination; 40% coursework and 60% examination

STAT7006   Design and analysis of sample surveys (6 credits)

Inferring the characteristics of a population from those observed in a selection or sample from that population is a situation often forced on us for economic, ethical or technological reasons. Against the background of practical situations, this course considers the basic principles, practice and design of sampling techniques to produce objective answers free from bias. Emphasis will be on current and local problems.
Assessment: One 3-hour written examination; 25% coursework and 75% examination

STAT7007   Categorical data analysis (6 credits)

Many social and medical studies, especially those involving questionnaires, contain large amounts of categorical data. Examples of categorical data include presence or absence of disease (yes / no), mode of transportation (bus, taxi, railway), attitude toward an issue (strongly disagree, disagree, agree, strongly agree). This course focuses on analyzing categorical response data with emphasis on hands-on training of analyzing real data using statistical software such as SAS. Consulting experience may be presented in the form of case studies. Topics include: classical treatments of 2 and 3-way contingency tables, measures of association and nonparametric methods; generalized linear models, logistic regression for binary, multinomial and ordinal data, loglinear models, Poisson regression; Modelling repeated measurements; generalized estimating equations.
Assessment: One 3-hour written examination; 50% coursework and 50% examination

STAT7008   Programming for data science (6 credits)

In the big data era, it is very easy to collect huge amounts of data. Capturing and exploiting the important information contained within such datasets poses a number of statistical challenges. This course aims to provide students with a strong foundation in computing skills necessary to use R or Python to tackle some of these challenges. Possible topics to be covered may include exploratory data analysis and visualization, collecting data from a variety of sources (e.g. Excel, web-scraping, APIs and others), object-oriented programming concepts and scientific computation tools. Students will learn to create their own R packages or Python libraries.
Assessment: 100% coursework

STAT7301   Socio-economic statistics for business and public policies (3 credits)

Huge volumes of socio-economic statistics are compiled and published on society and the economy by Governments and other bodies locally and elsewhere. Strong ability of business managers and authorities concerned to make effective reference to relevant data greatly enhances the quality of decision making in business and public policy processes. Students will learn about globally adopted standards for the compilation and dissemination of important data, such as those on population, labour, economic structure (in particular GDP), productivity, prices, trade, finance, housing, health and education; how to obtain them; and appropriate methods of utilizing them for the purposes of understanding socio-economic phenomena and making sound decisions. Ample practical examples drawn from Hong Kong and elsewhere will be presented.
Assessment: One 1.5-hour written examination; 40% coursework and 60% examination

STAT8000   Workshop on spreadsheet modelling and database management (3 credits)

This course aims to enhance students' IT knowledge and skills which are essential for career development of statistical and risk analysts. The course contains a series of computer hands-on workshops on Excel VBA programming, MS-Access and SQL and C++ basics.
Assessment: 100% coursework, assessment of this course is on a pass or fail or distinction basis

STAT8002   Project (6 credits)

A project in any branch of statistics or probability will be chosen under the supervision of individual staff member. A substantial written report is required. Availability of this course is subject to approval.
Pre-requisites: Students should not be taking or have taken STAT8089 Capstone Project or equivalent
Assessment: 60% written report and 40% oral presentation

STAT8003   Time series forecasting (6 credits)

A time series consists of a set of observations on a random variable taken over time. Such series arise naturally in climatology, economics, finance, environmental research and many other disciplines. In additional to statistical modelling, the course deals with the prediction of future behaviour of these time series. This course distinguishes different types of time series, investigates various representations for them and studies the relative merits of different forecasting procedures.
Assessment: One 3-hour written examination; 40% coursework and 60% examination

STAT8007   Statistical methods in economics and finance (6 credits)

This course provides a comprehensive introduction to state-of-the-art statistical techniques in economics and finance, with emphasis on their applications to time series and panel data sets in economics and finance. Topics include: regression with autocorrelated errors, modelling returns and volatility; instrumental variables and two stage least squares; panel time series models; unit root tests, co-integration, error correction models.
Assessment: One 3-hour written examination; 25% coursework and 75% examination

STAT8014   Risk management and Basel Accords (6 credits)

Being an important financial centre, Hong Kong has always been on the alert for risk in the banking and financial industry. We have weathered many attacks and crises over the past decades. Following the deep and long lasting global financial crisis started in 2007/08, this risk has been the primary focus of most people. This course will provide, and it is paramount for people in or related to the industry be fully aware of the relevant risk management, including the nature, the culture, the framework, the cycle, the measurement (with focus on market, credit and operational risks) and the mitigation techniques, along with the knowledge of the Basel Accords and practical critical issues.
Assessment: One 3-hour written examination; 40% coursework and 60% examination

STAT8015   Actuarial statistics (6 credits)

The main focus of this module will be on financial mathematics of compound interest with an introduction to life contingencies and statistical theory of risk. Topics include simple and compound interest, annuities certain, yield rates, survival models and life tables, population studies, life annuities, assurances and premiums, reserves, joint life and last survivor statuses, multiple decrement tables, expenses, individual and collective risk theory.
Assessment: One 3-hour written examination; 25% coursework and 75% examination

STAT8016   Biostatistics (6 credits)

Statistical methodologies and applications in fields of medicine, clinical research, epidemiology, biology and biomedical research are considered. The types of statistical problems encountered will be motivated by experimental data sets. Important topics include design and analysis of randomized clinical trials, group sequential designs and crossover trials; survival studies; diagnosis; statistical analysis of the medical process.
Assessment: One 2-hour written examination; 40% coursework and 60% examination

STAT8017   Data mining techniques (6 credits)

With the rapid developments in computer and data storage technologies, the fundamental paradigms of classical data analysis are mature for change. Data mining techniques aim at helping people to work smarter by revealing underlying structure and relationships in large amounts of data. This course takes a practical approach to introduce the new generation of data mining techniques and show how to use them to make better decisions. Topics include data preparation, feature selection, association rules, decision trees, bagging, random forests and gradient boosting, cluster analysis, neural networks, introduction to text mining.
Pre-requisites: Students should not be taking or have taken STAT8089 Capstone Project or equivalent
Assessment: 100% coursework

STAT8019   Marketing analytics (6 credits)

This course aims to introduce various statistical models and methodology used in marketing research. Special emphasis will be put on marketing analytics and statistical techniques for marketing decision making including market segmentation, market response models, consumer preference analysis and conjoint analysis. Contents include market response models, statistical methods for segmentation, targeting and positioning, statistical methods for new product design.
Assessment: One 3-hour written examination; 40% coursework and 60% examination

STAT8020   Quantitative strategies and algorithmic trading (6 credits)

Quantitative trading is a systematic investment approach that consists of identification of trading opportunities via statistical data analysis and implementation via computer algorithms. This course introduces various methodologies that are commonly employed in quantitative trading.

The first half of the course focuses at strategies and methodologies derived from the data snapshotted at daily or minute frequency. Some specific topics are: (1) techniques for trading trending and mean-reverting instruments, (2) statistical arbitrage and pairs trading, (3) detection of "time-series" mean reversion or stationarity, (4) cross-sectional momentum and contrarian strategies, (5) back-testing methodologies and corresponding performance measures, and (6) Kelly formula, money and risk management. The second half of the course discusses statistical models of high frequency data and related trading strategies. Topics that planned to be covered are: (7) introduction of market microstructure, (8) stylized features and models of high frequency transaction prices, (9) limit order book models, (10) optimal execution and smart order routing algorithms, and (11) regulation and compliance issues in algorithmic trading.
Pre-requisites: Pass in STAT6013 Financial data analysis or equivalent
Assessment: One 2-hour written examination; 60% coursework and 40% examination

STAT8021   Big data analytics (6 credits)

The recent explosion of social media and the computerization of every aspect of life resulted in the creation of volumes of mostly unstructured data (big data): web logs, e-mails, video, speech recordings, photographs, tweets and others. This course aims to provide students with knowledge and skills of some advanced analytics and statistical modelling for solving big data problems. Students are required to possess basic understanding of Python language.
Pre-requisites: Pass in STAT8017 Data mining techniques or equivalent; students should not be taking or have taken STAT8307 Text analytics or equivalent
Assessment: 100% coursework

STAT8088   Practicum (6 credits)

This course is open to students of Master of Statistics Programme only. It provides students with first-hand experience in the applications of academic knowledge in a real-life work environment. To be eligible, students should be undertaking a statistics-related or risk-management-related practicum with no less than 160 hours in at least 20 working days spent in a paid or unpaid position. It is possible for part-time students to complete their practicum within their current place of employment. The practicum will normally take place in the second semester or summer semester for full-time students or during the second year for part-time students.

Assessment: Upon completion of the practicum, each student is required to submit a written report and to give an oral presentation on his/her practicum experience. Supervisors will assess the students based on their performance during the practicum period. Assessment of this course is on a Pass or Fail or Distinction basis with 3 criteria: (1) supervisor's evaluation, (2) written report, (3) oral presentation. Please note that fail in fulfilling any of the 3 criteria satisfactorily would lead to a "Fail" grade in the course.

STAT8089   Capstone project (6 credits)

This project-based course aims to provide students with capstone experience to work on a real-world problem and carry out a substantial data analysis project which requires integration of the knowledge they have learnt in the curriculum. Students will work in small groups under the guidance of their supervisor(s). The project topic is not limited to academic context, but can also be extended to a community or corporate outreach project. Students will need to find an interesting topic of their own, conduct literature search regarding the most recent research related to the problem, make suggestions to improve the current situations or even solve the problem identified in their project. A substantial written report is required.
Pre-requisites: Students should not be taking or have taken STAT8002 Project or STAT8017 Data mining techniques or equivalent
Assessment: 15% project proposal; 50% written report and 35% oral presentation

---

STAT8300   Career development and communication workshop (3 credits)

The course is specially designed for students who wish to sharpen their communication and career preparation skills through a variety of activities including lectures, skill-based workshops, small group discussion and role plays. All of which aim to facilitate students in making informed career choices, provide practical training to enrich communication, presentation, time management and advanced interview skills, and to enhance students' overall competitiveness in the employment markets.
Assessment: 100% coursework, assessment of this course is on a pass or fail or distinction basis

---

STAT8302   Structural equation modelling (3 credits)

Structural Equation Modelling (SEM) is a general statistical modelling technique to establish relationships among variables. A key feature of SEM is that observed variables are understood to represent a small number of "latent constructs" that cannot be directly measured, only inferred from the observed measured variables. This course covers the theories of structural equation models and their applications. Topics may include path models, confirmatory factor analysis, structural equation models with latent variables, Sub-models including multiple group analysis, MIMIC model, second order factor analysis, two-wave model, and simplex model, model fitness, model identification, and Comparison with competing models.
Pre-requisites: Pass in STAT7005 Multivariate methods or equivalent
Assessment: One 1.5-hour written examination; 50% coursework and 50% examination

STAT8304   Current topics in Statistics (3 credits)

The purpose of this course is to broaden the students' knowledge of statistics by studying some contemporary topics motivated by applications of statistics. These topics will build on the theory and methods covered in the compulsory courses. The topics offered each year depend on student interests and staff availability. After completing the course, students will acquire knowledge and skills of some advanced statistical techniques for solving real life problems.

Assessment: One 1.5-hour written examination; 40% coursework and 60% examination

---

STAT8305   Bayesian statistics (3 credits)

This course introduces Bayesian methodologies and computational techniques of Markov chain Monte Carlo (MCMC). It covers fundamental Bayesian concepts, modeling and inference, including prior specification, posterior distribution, posterior predictive, Bayes factor, Bayesian hypothesis testing, Bayesian hierarchical modeling, and Bayesian decision theoretic analysis. From the computational perspective, it covers rejection sampling, importance sampling, Metropolis-Hastings algorithm, Gibbs sampling, and data augmentation MCMC techniques. Statistical software R and Python will be used for Bayesian computation.
Pre-requisites: Students should not be taking or have taken STAT6011 Computational statistics or equivalent
Assessment: One 1-hour written examination; 25% coursework and 75% examination

---

STAT8306   Statistical methods for network data (3 credits)

The six degrees of separation theorizes that human interactions could be easily represented in the form of a network. Examples of networks include router networks, the World Wide Web, social networks (e.g. Facebook or Twitter), genetic interaction networks and various collaboration networks (e.g. movie actor coloration network and scientific paper collaboration network). Despite the diversity in the nature of sources, the networks exhibit some common properties. For example, both the spread of disease in a population and the spread of rumors in a social network are in sub-logarithmic time. This course aims at discussing the common properties of real networks and the recent development of statistical network models. Topics may include common network measures, community detection in graphs, preferential attachment random network models, exponential random graph models, models based on random point processes and the hidden network discovery on a set of dependent random variables.
Assessment: One 1.5-hour written examination; 50% coursework and 50% examination

---

STAT8307   Text analytics (3 credits)

The textual data constitutes and enormous proportion of unstructured data which is characterized as one of 'V's in Big Data. The logical and computational reasonings are applied to transform large collection of written resources to structured data for use in further analysis, visualization, integration with structured data in database or warehouse, and further refinement using machine learning systems. This course introduces the methodology of text mining and text analytics. Topics are selected from natural language processing, word representation, text categorization and clustering, topic modelling and sentiment analysis. Students are required to possess basic understanding of Python language.
Pre-requisites: Pass in STAT8017 Data mining techniques or equivalent; Students should not be taking or have taken STAT8021 Big data analytics or equivalent
Assessment: 100% coursework