



**FAKULTI SAINS**  
*Faculty of Science*

**Institute of Mathematical Sciences**

**PROGRAMME HANDBOOK**

**Master of Science in Statistics  
by Coursework**

**ACADEMIC SESSION 2025/2026**

<https://ism.um.edu.my/postgraduate-info>

| <b>Master of Science in Statistics<br/>2025/2026 Academic Session<br/>(42 CREDITS)</b> |   |                |
|--|---|----------------|
| <b>1. Programme Core Courses (27 CREDITS)</b>  |   |                |
| <b>Course Code</b>   | <b>Course Name</b>                          | <b>Credits</b> |
| SQB7022  | Research Methodology for Statistics         | 3              |
| SQB7024  | Statistical Inference                       | 4              |
| SQB7025  | Probability Theory                          | 4              |
| SQB7026  | Programming in Statistics                   | 3              |
| SQB7027  | Statistical Consultancy and Data Analysis   | 3              |
| SQB7038  | Research Project in Statistics              | 10             |
| <b>2. Programme Elective Courses (15 CREDITS)</b>                                      |   |                |
| <b>Course Code</b>   | <b>Course Name</b>                          | <b>Credits</b> |
| SQB7028  | Multivariate Analysis                       | 3              |
| SQB7029  | Stochastic Models                           | 3              |
| SQB7030  | Bayesian Statistics                         | 3              |
| SQB7031  | Generalized Linear Models                   | 3              |
| SQB7032  | Experimental Design and Quality Engineering | 3              |
| SQB7033  | Statistical Time Series                     | 3              |
| SQB7034  | Computer Intensive Methods                  | 3              |
| SQB7035  | Robust Statistics                           | 3              |
| SQB7036  | Data Mining                                 | 3              |
| SQB7037  | Survival Data Analysis                      | 3              |

## **PROGRAMME GOAL**

To produce master graduates who have strong statistical knowledge, are capable of analysing and solving problems, and can think critically. Our graduates are capable of adapting to diverse environments and contributing meaningfully towards professions in different fields.

## **PROGRAMME LEARNING OUTCOMES**

At the end of the programme, graduates of the Master of Science in Statistics are expected to demonstrate the following qualities:

1. Comprehend advanced statistical theories with regards to the statistical and mathematical arguments, proofs and abstract concepts.
2. Apply technical and practical skills related to advanced statistics in different fields.
3. Perform professional work with good social skills, and behaving responsibly to the community.
4. Practise professionally and ethically in one's profession.
5. Work collaboratively, and articulate statistical ideas clearly, accurately, and effectively.
6. Solve statistical problems using logical, analytical, and critical thinking.
7. Habitually use statistical thinking to understand and manage information.

## SAMPLE PLANNER FOR COURSES (GRADUATION IN TWO SEMESTERS)

| COMPONENT        | YEAR 1   |         |  |         | TOTAL CREDITS |
|------------------|--|---------|--|---------|---------------|
|                  | SEMESTER 1                                     |         | SEMESTER 2   |         |               |
|                  | COURSES  | CREDITS | COURSES  | CREDITS |               |
| Core Courses     | SQB7022<br>Research Methodology for Statistics | 3       | SQB7024<br>Statistical Inference                     | 4       | 27            |
|                  | SQB7025<br>Probability Theory                  | 4       | SQB7027<br>Statistical Consultancy and Data Analysis | 3       |               |
|                  | SQB7026<br>Programming in Statistics           | 3       | SQB7038<br>Research Project in Statistics            | 5       |               |
|                  | SQB7038<br>Research Project in Statistics      | 5       |  |         |               |
| Elective Courses | SQB70**  | 6       | SQB70**  | 9       | 15            |
| TOTAL CREDITS    |  | 21      |  | 21      | 42            |

**Note:** Most of our programme alumni take 3 to 4 semesters to finish their studies. Please consult the course coordinator about course planning.

## INSTITUTE OF MATHEMATICAL SCIENCES

The Institute of Mathematical Sciences (ISM) was established as a department in the Faculty of Science when the University of Malaya was founded in Kuala Lumpur in 1959. At present, the Institute offers four undergraduate degree programmes, one master's by coursework programme and master's by research and PhD programmes.

Since the 2017/2018 academic session, ISM has been conducting the Master of Science in Statistics coursework programme. This programme provides opportunities to first degree graduates to acquire advanced statistical knowledge and skills.

### STAFF

ISM has a group of experienced lecturers in teaching. They are active in research and have published their work in international as well as local journals. Our research activities encompass a broad spectrum, from findings and knowledge which are abstract in nature, to those with direct applications in the industry. ISM strives to establish and forge close relationships with the industry and other research institutions. This strengthens the quality of teaching and supervision of projects/theses for students at the bachelor, master and doctoral levels.

#### HEAD:

**Professor Dr. Ibrahim Mohamed**, *BSc(Bristol), MSc(Reading), PhD(UiTM)*

#### DEPUTY HEADS:

**Professor Dr. Ng Kok Haur**, *BSc, MSc(UPM), PhD(UM)*

**Assoc. Prof. Dr. Adriana Irawati Nur Ibrahim**, *BSc(USM), MSc(UM), PhD(Bath)*

**Assoc. Prof. Raveendran VGK Menon**, *BEng(UM), MActuarialSc(Georgia State), ASA, AIAA*

**Dr. Muhamad Hifzhudin Noor Aziz**, *BSc(UM), MSc, PhD(Glasgow)*

**Dr. Siti Suzlin Supadi**, *BSc, MSc, PhD(UM)*

### MATHEMATICS UNIT

#### COORDINATOR (B.Sc. in MATHEMATICS):

**Dr. Wang Kah Lun**, *BSc, MSc(UKM), PhD(UM)*

#### PROFESSORS:

**Dr. Wong Kok Bin**, *BSc, MSc, PhD(UM)*

**Dr. Zailan Siri**, *BSc(UM), MSc(UPM), PhD(UKM)*

#### HONORARY PROFESSORS:

**Dr. Kurunathan Ratnavelu**, *BSc, MSc(UM), PhD(Flinders)*

**Dr. Mohd Omar**, *BSc(UM), MSc(Hull), PhD(Exeter)*

**Dr. Wan Ainun Mior Othman**, *BSc(UNCC), MSc(N Carolina State), PhD(USM)*

#### ASSOCIATE PROFESSORS:

**Dr. Chooi Wai Leong**, *BSc, MSc, PhD(UM)*

**Dr. Deng Chai Ling**, *BSc, MSc, PhD(UM)*

**Dr. Noor Fadiya Mohd Noor**, *BSc, MSc(UTM), PhD(UKM)*

#### SENIOR LECTURERS:

**Dr. Amizah Malip**, *BSc(IIUM), MSc, PhD(London)*

**Dr. Avik De**, *MSc, PhD(Calcutta)*

**Dr. Chng Zhi Yee**, *BSc(UM), PhD(UNSW)*  
**Dr. Kohilavani Naganthran**, *BSc(UNISEL), MSc, PhD(UKM)*  
**Dr. Kwa Kiam Heong**, *BSc, MSc(UM), PhD(Ohio)*  
**Dr. Loo Tee How**, *BSc, MSc, PhD(UM)*  
**Dr. Lau Jin Ting**, *BSc, MSc, PhD(UM)*  
**Dr. Mohd Zahurin Mohamed Kamali**, *BSc, MSc, PhD(UM)*  
**Dr. Muhamad Hifzhudin Noor Aziz**, *BSc(UM), MSc, PhD(Glasgow)*  
**Dr. Nur Fadhilah Mohd Shari**, *BSc(Purdue), MSc, PhD(UM)*  
**Dr. Ong Siew Hui**, *BSc, MSc, PhD(UM)*  
**Dr. Oon Shea Ming**, *BSc, MSc, PhD(UHP)*  
**Dr. Qua Kiat Tat**, *BSc(UM), MSc, PhD(UM)*  
**Dr. Ruhaila Md. Kasmani**, *BSc(UKM), MSc(UTM), PhD(UM)*  
**Dr. Shahizat Amir**, *BSc(UKM), MPhil, PhD(UM)*  
**Dr. Siti Suzlin Supadi**, *BSc, MSc, PhD(UM)*  
**Dr. Tan Ta Sheng**, *BA, MA, PhD(Cambridge)*  
**Dr. Wang Kah Lun**, *BSc, MSc(UKM), PhD(UM)*

## STATISTICS AND ACTUARIAL SCIENCE UNIT

### COORDINATOR (B.Sc. in STATISTICS):

Assoc. Prof. Dr. Rossita Mohamad Yunus, *BSc, MSc(UM), PhD(USQ)*

### COORDINATOR (B. ACTUARIAL SCIENCE):

Dr. Syaza Nawwarah Zein Isma, *BSc(UM), MSc(IIUM), PhD(ANU)*

### COORDINATOR (M.Sc. in STATISTICS):

Assoc. Prof. Dr. Khang Tsung Fei, *BSc, MSc(UM), PhD(NUS)*

### PROFESSORS:

Dr. Ibrahim Mohamed, *BSc(Bristol), MSc(Reading), PhD(UiTM)*

Dr. Ng Kok Haur, *BSc, MSc(UPM), PhD(UM)*

### HONORARY PROFESSORS:

Dr. Nor Aishah Hamzah, *BSc(Southampton), MSc(Leeds), PhD(Bristol), DipEd(UKM), MIS(UK)*

Dr. Ong Seng Huat, *BSc, MSc, PhD(UM)*

### ASSOCIATE PROFESSORS:

Dr. Adriana Irawati Nur Ibrahim, *BSc(USM), MSc(UM), PhD(Bath)*

Dr. Khang Tsung Fei, *BSc, MSc(UM), PhD(NUS)*

Mr. Raveendran A/L VGK Menon, *BEng(UM), MActuarialSc(Georgia State), ASA, AIAA*

Dr. Rossita Mohamad Yunus, *BSc, MSc(UM), PhD(USQ)*

### SENIOR LECTURERS:

Dr. Adina Najwa Kamarudin, *BSc, MSc (UiTM), PhD(Liverpool)*

Dr. Aneesha Pillay A/P Balachandran Pillay, *BSc, MSc, PhD(UM)*

Dr. Dharini A/P Pathmanathan, *BSc, MSc, PhD(UM)*

Dr. Koh You Beng, *BSc(UMS), MSc(UM), PhD(HKU)*

Dr. Mohd Azmi Haron, *BBA, MBA, PhD(UPM)*

Mdm. Nadiah Zabri, *BBA(Wisconsin-Madison), MSc(Kent), AIA*

Dr. Ng Choung Min, *BSc(UTM), MSc, PhD(UM)*

Dr. Nur Anisah Mohamed @ A. Rahman, *BSc, MSc(UM), PhD(Newcastle)*

Dr. Shaiful Anuar Abu Bakar, *DipActuarialSc, BSc(UiTM), MSc(Heriot-Watt), PhD(Manchester)*

Dr. Siti Fatimah Ismail, *BSc(UiTM), MSc, PhD(USM)*

Dr. Syaza Nawwarah Zein Isma, *BSc(UM), MSc(IIUM), PhD(ANU)*

Dr. Tan Shay Kee, *BSc, MSc, PhD(UM)*

### COORDINATOR (B.Sc. Ed. Mathematics):

Dr. Shahizat Amir, *BSc(UKM), MPhil, PhD(UM)*

## RESEARCH AREAS

Research areas at ISM include the following fields: differential geometry, group theory, ring theory, linear preserver problems, functional identities, linear and multilinear algebra, matrix theory, combinatorial graph theory, graph theory, social network analysis, supply chain management, operations research, numerical analysis, computational statistics, robust statistics, probability distribution theory, nonlinear time series, image processing, regression analysis, statistical quality control, mathematical biology, and data science.

## COMPUTER FACILITIES

Currently, ISM has a computer lab equipped with 10 laptops, 17 workstations, 121 desktops, 3 laser printers, 1 colour printer, and 4 heavy-duty dot matrix printers, all of which are interconnected through a network system. The lab is also equipped with 4 LCD projectors, 2 visualizers, and 3 scanners. The lab utilises state-of-the-art software such as MATLAB (with various toolboxes), SPSS, Wolfram Mathematica, MathType, Minitab, Microsoft Visual C++, Dev-C++, and S-PLUS. In addition, three of the lecture halls and tutorial rooms are each equipped with an LCD projector and a visualiser.

## JOB OPPORTUNITIES

The learning of mathematics increases one's skills in problem solving and analysis. It trains one's mind to process information, form sophisticated ideas, and discern complex arguments. The training to think quantitatively, logically, and analytically is invaluable to one's career.

Since the use of mathematics is all encompassing in human endeavours, a graduate's career opportunity is numerous and not confined to only teaching and research. Our graduates are well-represented in diverse sectors such as finance, business, industry, and information science.

## SYNOPSIS OF COURSES

### SQB7022 RESEARCH METHODOLOGY FOR STATISTICS

This course is designed to give knowledge and skills to students related to suitable methodologies in statistical research. It includes searching and critically evaluating journal articles on statistical problems. The course will introduce fundamental statistical concepts and techniques useful for research in statistics. The students will be guided in writing research proposal in the area of statistics.

#### Assessment

Continuous Assessment: 100%

#### References:

1. Agresti, A. (2015). *Foundations of linear and generalized linear models*. John Wiley & Sons.
2. Bordens, K.S., & Bruce, B. (2011). *Research and design methods: A process approach* (8th ed.). McGraw-Hill.
3. Flick, U. (2011). *Introducing research methodology: A beginner's guide to doing a research project*. Sage Publication.
4. Hogg, R.V., McKean, J.W., & Craig, A.T. (2013). *Introduction to mathematical statistics* (7th ed.). Pearson.



## SQB7024 STATISTICAL INFERENCE

Estimating of parameters and fitting of probability distribution; EM algorithm; use of statistical software for parameter estimation; Common families of distribution: location and scale families, exponential family; Principles of data reduction: the sufficiency principle – sufficient statistics, minimal sufficient statistics, ancillary statistics, complete sufficient statistics; Hypothesis tests: likelihood ratio tests, generalized likelihood ratio tests, error probabilities and power function; uniformly most powerful (UMP) test; UMP unbiased test.

### Assessment:

Continuous Assessment: 60%  
Final Examination: 40%

### References:

1. Casella, G., & Berger, R.L. (2008). *Statistical inference*. Wadsworth & Brooks/Cole.
2. Devore, J.L., & Berk, K.N. (2012). *Modern mathematical statistics with applications*. Cengage Learning.
3. Hogg, R.V., Craig, A.T., & McKean, J.W. (2004). *Introduction to mathematical statistics* (6th ed.). Collier MacMillan.
4. Mukhopadhyay, N. (2020). *Probability of statistical inference*. CRC Press.

## SQB7025 PROBABILITY THEORY

Probability measure and space, sigma field. Lebesgue integration. Random variables, measurability, independence. Distribution functions. Inequalities, characteristic functions. Various modes of convergence of sequences of random variables. Classical limit theorems. Examples of applications.

### Assessment:

Continuous Assessment: 60%  
Final Examination: 40%

### References:

1. Bhattacharya, R., & Waymire, E.C. (2016). *A basic course on probability theory*. Springer.
2. Borovkov, A. A. (2013). *Probability theory*. Springer.
3. Karr, A. F. (1993). *Probability*. Springer-Verlag.
4. Ross, S. M. (2019). *Introduction to probability models* (12th ed.). Academic Press.
5. Rotar, V. I. (1997). *Probability theory*. World Scientific.

## SQB7026 PROGRAMMING IN STATISTICS

Use of functions and commands in statistical packages for exploratory data analysis, modelling and statistical inferences, coding programming language.

### Assessment

Continuous Assessment: 60%  
Final Examination: 40%

### References:

1. Chambers, J.M., Cleveland, W.S., Kleiner, B., & Tukey, P.A. (1983). *Graphical methods for data analysis*. Chapman and Hall.
2. Cornillon, P.-A. et al. (2012). *R for statistics*. CRC Press.
3. Esbensen, K.H., Guyot, D., Westad, F., & Houmoller, L.P. (2009). *Multivariate data analysis: In practice*. CAMO.

4. Lander, J.P. (2017). *R for everyone: Advanced analytics and graphics* (2nd ed.). Addison-Wesley.
5. Venables, W.N., & Ripley, B.D. (2002). *Modern applied statistics with S*. Springer-Verlag.

## **SQB7027 STATISTICAL CONSULTANCY AND DATA ANALYSIS**

Introduction to consultancy activities and consulting methods. Related problems and issues. Exposure to the use of primary and secondary data from various sources. Application of suitable statistical methods such as multivariate analysis, regression and time series in the analysis of real data. Producing report and presenting the findings that suit the needs of the clients.

### Assessment

Continuous Assessment: 100%

### References:

1. Boen, J.R., & Zahn, D.A. (1982). *The human side of statistical consulting*. Lifetime Learning Publications.
2. Hand, D.J., & Everitt, B.S. (1987). *The statistical consultant in action*. Cambridge University Press.
3. Tufte, E.R. (2001). *The visual display of quantitative information*. (2nd ed.). Graphics Press.
4. Wickham, H., & Grolemund, G. (2017). *R for data science: Import, tidy, transform, visualize, and model data*. O'Reilly Media.

## **SQB7028 MULTIVARIATE ANALYSIS**

Graphical display of multivariate data. Dimensional reduction methods (factor analysis, principal components, cluster analysis, etc.). Discriminant and classification methods (discriminant analysis and cluster analysis). Multivariate linear regression, canonical correlation analysis. Selected topics in advanced multivariate methods (e.g. sparse multivariate methods, multidimensional scaling, and functional data analysis). Introduction to applied spatial statistics. Exploratory spatial data analysis. Selected topics in spatial data analysis (spatial regression, random field, etc.). Models and methods for spatiotemporal and multivariate spatial data. Application of multivariate analysis to real world problems.

### Assessment:

Continuous Assessment: 60%

Final Examination: 40%

### References:

1. Anderson, T. W. (1984). *An introduction to multivariate statistical analysis* (2nd ed.). John Wiley.
2. Arnold, S. F. (1981). *The theory of linear models and multivariate statistical analysis*. John Wiley.
3. Chatfield, C., & Collins, A. J. (1980). *Introduction to multivariate analysis*. Chapman & Hall.
4. Konishi, S. (2014). *Introduction to multivariate analysis - Linear and nonlinear modelling*. Chapman & Hall.
5. Johnson R.A., & Wichern, D.W. (2007). *Applied multivariate statistical analysis* (6th ed.). Pearson International Education.

## **SQB7029 STOCHASTIC MODELS**

Introduction to Markov chains. Continuous-time Markov chains. Poisson processes. Backward and forward Kolmogorov equations. Birth and death processes with examples. Definition and concepts in renewal processes, distribution for the number of renewal, renewal function and theorems for renewal processes. Examples for various types and applications of renewal processes. Introduction to martingales. Examples of stochastic models for real world applications.

Assessment:

Continuous Assessment: 60%  
Final Examination: 40%

References:

1. Cox, D. R. (1962). *Renewal theory*. Methuen and Company, Ltd.
2. Cox, D. R., & Miller, H. D. (1977). *The theory of stochastic processes*. Chapman & Hall.
3. Durrett, R. (2016). *Essentials of stochastic processes*. Springer.
4. Taylor, H. M., & Karlin, S. (1998). *An introduction to stochastic modelling*. Academic Press.

## **SQB7030 BAYESIAN STATISTICS**

Bayesian inference. Prior formulation. Multi-parameter models. Implementation via posterior sampling. Convergence assessment. Hierarchical and mixture models. Model checking. Model evaluation and comparison. Bayesian decision theory. Examples with different models from different areas of real-world application such as Bayesian linear regression and Bayesian network.

Assessment:

Continuous Assessment: 60%  
Final Examination: 40%

References:

1. Berger, J. O. (2010). *Statistical decision theory and Bayesian analysis* (2nd ed.). Springer-Verlag.
2. Cover, T.M., & Thomas, J.A. (2006). *Elements of information theory* (2nd ed.). John Wiley.
3. Cowles, M. K. (2013). *Applied Bayesian statistics with R and Open Bugs example*. Springer.
4. Gelman, A. (1995). *Bayesian data analysis*. Chapman & Hall.
5. Lee, P.M. (1997). *Bayesian statistics: An introduction*. Arnold.
6. Polson, N.G., & Tiao, G.C. (1995). *Bayesian inference*. Edward Elgar Pub.

## **SQB7031 GENERALIZED LINEAR MODELS**

Generalised Linear Models (GLM) based on exponential family. Linear predictor, link function, canonical link, likelihood, the iterative reweighted least squares algorithm, Fisher information, tests on individual parameters, deviance, residuals. Application of GLMs on real life data such as using models for continuous and discrete response, models for polytomous data, log-linear models, quasi-likelihoods, models for correlated responses, and generalised additive model. Statistical software use for data analysis. Model fitting, selection and diagnostics.

Assessment:

Continuous Assessment: 60%  
Final Examination: 40%

References:

1. Agresti, A. (2013). *Categorical data analysis* (3rd ed.). Wiley-Interscience.
2. Dobson A.J., & Barnett, A.G. (2008). *An introduction to generalized linear models* (3rd ed.). Chapman & Hall/CRC.
3. McCullagh, P., & Nelder J.A. (1989). *Generalized linear models* (2nd ed.). Chapman & Hall.
4. Myers, R.H., Montgomery, D.C., Vining, G.G., & Robinson, T.J. (2010). *Generalized linear models: With applications in engineering and the sciences* (2nd ed.). John Wiley & Sons.
5. Smithson, M., & Merkle, E.C. (2014). *Generalized linear models for categorical and continuous limited dependent variables*. Chapman & Hall/CRC.

## **SQB7032 EXPERIMENTAL DESIGN AND QUALITY ENGINEERING**

Factorial designs. Blocking and confounding systems. Fractional factorial designs. Response surface methodology. Robust design. Gauge repeatability and reproducibility.

### Assessment:

Continuous Assessment: 60%  
Final Examination: 40%

### References:

1. Box, G.E.P., Hunter, J.S., & Hunter, W.G. (2005). *Statistics for experimenters*. John Wiley.
2. Fowlkes, W.Y., & Creveling, C.M. (1995). *Engineering methods for robust product design: Using Taguchi methods in technology and product development*. Addison-Wesley.
3. Montgomery, D.C. (2012). *Design and analysis of experiments* (8th ed.). John Wiley.
4. Myers, R.H., Montgomery, D.C., & Anderson-Cook, C. M. (2009). *Response surface methodology: Process and product using designed experiments* (3rd ed.). John Wiley.
5. Wu, C.F.J., & Hanada, M.S. (2009). *Experiments: Planning, analysis and optimization* (2nd ed.). John Wiley.

## **SQB7033 STATISTICAL TIME SERIES**

Introduction to time series and forecasting, time series regression, ARIMA models, state space models, dynamic regression models.

### Assessment:

Continuous Assessment: 60%  
Final Examination: 40%

### References:

1. Box, G.E.P., Jenkins, G.W., & Reinsel, G. (2008) *Time series analysis, forecasting and control* (4th ed.). Prentice Hall.
2. Brockwell, P. J., & Davis, R. A. (2010). *Introduction to time series and forecasting* (2nd ed.). Springer.
3. Cryer, J.D. (2008). *Time series analysis: With application in R*. Springer.
4. Shumway, R. H., & Stoffer, D. S. (2011). *Time series analysis and its applications: With R examples* (2nd ed.). Springer.

## **SQB7034 COMPUTER INTENSIVE METHODS**

Error in floating point calculations. Probability function and distribution function approximations. Generating random numbers, including evaluating the quality of the generator and calculation methods in linear algebra: Gaussian elimination, sweep operators. Calculation methods for multiple regression, (not constrained) nonlinear regression and model fitting other than the least squares, bootstrap, Markov chain Monte Carlo and EM algorithm.

### Assessment:

Continuous Assessment: 60%  
Final Examination: 40%

### References:

1. Chernick, M.R. (2011). *Bootstrap methods: A guide for practitioners and researchers* (2nd ed.). Wiley-Interscience.

2. Gentle, J.E., Hardle, W.K., & Mori, Y. (2012). *Handbook of computational statistics: Concept and methods*. Springer.
3. Hjorth, J.S. (1994). *Computer intensive statistical methods: Validation, model selection and bootstrap*. Chapman & Hall.
4. Kennedy, W.J., & Gentle, J.E. (1980). *Statistical computing*. Marcel Dekker.
5. Peng, R.D. (2016). *R programming for data science*. Leanpub.

## **SQB7035 ROBUST STATISTICS**

Introduction to robust statistics: a review of basic philosophy and concepts central to the area of robust statistics, as well as a review of some basic robust methods. Topics covered include the L-estimates, monotonic and the redescending M-estimates, with applications to the estimation of univariate location and scale; robust regression methods including the M-estimates, the GM-estimates, the least median of squares method (LMS), the S-estimates and the MM-estimates; robust multivariate methods including the M-estimates, the minimum volume ellipsoid method (MVE), and the MM-estimates. Notions of robustness such as breakdown point and influence function will also be reviewed. Computational issues and analysis of real data sets using R software will also be discussed.

### Assessment:

Continuous Assessment: 60%  
Final Examination: 40%

### References:

1. Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., & Stahel, W. A. (1986). *Robust statistics: The approach based on influence functions*. Wiley.
2. Huber, P. J., & Ronchetti, E.M. (2009). *Robust statistics* (2nd ed.). John Wiley & Sons.
3. Maronna, R.A., Martin, D.R., & Yohai, V.J. (2006). *Robust statistics: Theory and methods*. John Wiley & Sons.
4. Rousseeuw, P.J., & Leroy, A. M. (2003). *Robust regression and outlier detection*. Wiley-Interscience.

## **SQB7036 DATA MINING**

Introduction to data mining methods and tools. Data cleaning and processing. Data visualisation. Dimensional reduction methods: principal component analysis, variable selection. Data mining algorithms for supervised learning (logistic regression, support vector machine, discriminant analysis, classification and regression trees, etc.) and unsupervised learning (K-means, hierarchical clustering). Ensemble methods (random forests, bagging, boosting). Performance evaluation of classifiers. Interpretation of data mining results. Examples of data mining for real world applications.

### Assessment:

Continuous Assessment: 60%  
Final Examination: 40%

### References:

1. Bramer, M. (2013). *Principles of data mining* (2nd ed.). Springer-Verlag.
2. Flach, P. (2012). *Machine learning: The art and science of algorithms that make sense of data*. Cambridge University Press.
3. Hastie, T., Tibshirani, R., & Friedman, J. (2017). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.
4. Wickham, H., & Grolemund, G. (2017). *R for data science: Import, tidy, transform, visualize, and model data*. O'Reilly Media.

### **SQB7037 SURVIVAL DATA ANALYSIS**

Basic concepts such as survival and hazard functions. Survival data analysis including life table, Kaplan-Maier; log-rank and Wilcoxon tests. Survival regression modeling including the Cox regression model, several parametric models and the accelerated life time model and risk model. Diagnostic checking of the models. Application to the real dataset.

#### Assessment

Continuous Assessment: 60%

Final Examination: 40%

#### References:

1. Collett, D. (2015). *Modelling survival data in medical research* (3rd ed.). Chapman and Hall
2. Hosmer, D.W., Lemeshow, S., & May, S. (2011). *Applied survival analysis: Regression modeling of time to event data*. John Wiley.
3. Lee, E.T. (2013). *Statistical methods for survival data analysis* (2nd ed.). John Wiley & Sons.
4. Miller, R.G. (2011). *Survival analysis* (2nd ed.). John Wiley & Sons.

### **SQB7038 RESEARCH PROJECT IN STATISTICS**

Refer to lecturers concerned for project synopsis and reference texts.

#### Assessment

Continuous Assessment: 100%

#### References:

Refer to lecturers.