City University of Hong Kong
Course Syllabus

offered by School of Data Science
with effect from Semester A 2019/20

Part I Course Overview

Course Title: **Computational Statistics**

Course Code: **SDSC3005**

Course Duration: **One Semester**

Credit Units: **3**

Level: **B3**

Proposed Area:
(for GE courses only)

- ☐ Arts and Humanities
- ☐ Study of Societies, Social and Business Organisations
- ☐ Science and Technology

Medium of Instruction: **English**

Medium of Assessment: **English**

Prerequisites:
(Course Code and Title) **SDSC3007 Advanced Statistics**

Precursors:
(Course Code and Title) **Nil**

Equivalent Courses:
(Course Code and Title) **Nil**

Exclusive Courses:
(Course Code and Title) **Nil**

Part II Course Details

1. Abstract

(A 150-word description about the course)

This course introduces students to algorithms and techniques for statistical computing and their implementations through R software. Students will learn important computational statistics methods such as the EM algorithm, Fisher's scoring, Monte Carlo simulation, Markov chain Monte Carlo, and bootstrap. Additionally, students will learn statistical applications of these methods, the key advantages of using each method, and how they can be coded in R. Efficient programming methods for R will be taught. Therefore, students gain knowledge of many different tools that can be combined to solve statistical computing problems. Assignments will involve the use R.

2. Course Intended Learning Outcomes (CILOs)

(CILOs state what the student is expected to be able to do at the end of the course according to a given standard of performance.)

No.	CILOs#	Weighting* (if applicable)	Discovery-enriched curriculum related learning outcomes (please tick where appropriate)		
			A1	*A2*	*A3*
1.	**Understand** the application background of statistical computing algorithm and techniques.	10%	√		
2.	**Elaborate** the theories behind the algorithms in computational statistics.	20%	√	√	
3.	**Implement** various types of statistical computing algorithms in R.	20%		√	√
4.	**Apply** the correct algorithm to solve a statistical computing problem.	30%	√	√	√
5.	**Discuss** tuning parameters for various statistical computing algorithms.	20%	√	√	
		100%			

** If weighting is assigned to CILOs, they should add up to 100%.*

 # Please specify the alignment of CILOs to the Gateway Education Programme Intended Learning outcomes (PILOs) in Section A of Annex.

A1: Attitude
Develop an attitude of discovery/innovation/creativity, as demonstrated by students possessing a strong sense of curiosity, asking questions actively, challenging assumptions or engaging in inquiry together with teachers.

A2: Ability
Develop the ability/skill needed to discover/innovate/create, as demonstrated by students possessing critical thinking skills to assess ideas, acquiring research skills, synthesizing knowledge across disciplines or applying academic knowledge to self-life problems.

A3: *Accomplishments*
Demonstrate accomplishment of discovery/innovation/creativity through producing /constructing creative works/new artefacts, effective solutions to real-life problems or new processes.

3. Teaching and Learning Activities (TLAs)

(TLAs designed to facilitate students' achievement of the CILOs.)

TLA	Brief Description	CILO No.					Hours/week (if applicable)
		1	2	3	4	5	
Lectures	Learning through teaching is primarily based on lectures. Mini-lectures and small-group exercises will be used to facilitate conceptual understanding and applications of various statistical tools and techniques.	√	√	√	√	√	26 hours/semester
Tutorial Exercises	The team-based exercises provide students with the opportunities to familiarize and apply the statistical tools learnt during the lectures through practical problem solving.			√	√	√	13 hours/semester

4. Assessment Tasks/Activities (ATs)

(ATs are designed to assess how well the students achieve the CILOs.)

Assessment Tasks/Activities	CILO No.					Weighting*	Remarks
	1	2	3	4	5		
Continuous Assessment: 60%							
Test		√	√	√		30%	
Assignments	√	√	√	√	√	30%	
Examination: 40% (duration: 2 hours)							
Examination	√	√	√	√	√	40%	
						100%	

** The weightings should add up to 100%.*

For a student to pass the course, at least 30% of the maximum mark for the examination should be obtained

5. Assessment Rubrics

(Grading of student achievements is based on student performance in assessment tasks/activities with the following rubrics.)

Assessment Task	Criterion	Excellent (A+, A, A-)	Good (B+, B, B-)	Adequate (C+, C, C-)	Marginal (D)	Failure (F)
1. Test	2-hour test to assess students' understanding of computational statistics methods and algorithms.	High	Significant	Moderate	Basic	Not even reaching marginal levels
2. Assignments	Students' ability to correctly apply computational statistics methods in R to solve given statistics problems.	High	Significant	Moderate	Basic	Not even reaching marginal levels
3. Examination	Examination questions are designed to assess student's level of achievement of the intended learning outcomes, with emphasis placed on understanding and correct application, mostly through mathematical exposition, clear explanation, and numerical calculation, of the various computational statistics techniques.	High	Significant	Moderate	Basic	Not even reaching marginal levels

The test and assignments will be numerically-marked, while examination will be numerically-marked and grades-awarded accordingly.

Part III Other Information (more details can be provided separately in the teaching plan)

1. Keyword Syllabus

(An indication of the key topics of the course.)

Introduction to R

- Programming in R

Monte Carlo simulation

- Random number generation: inverse transform method, rejection sampling, ratio-of-uniforms method
- Monte Carlo methods, importance sampling

Computation of maximum likelihood estimate

- Fisher scoring and Newton's method
- EM algorithm

Computationally intensive frequentist inference methods

- Jackknife
- Bootstrap methods
- Cross-validation

Bayesian computation

- Metropolis-Hastings algorithm
- Gibbs sampling
- Sliced sampling
- Reversible jump Markov chain Monte Carlo

2. Reading List

2.1 Compulsory Readings

(Compulsory readings can include books, book chapters, or journal/magazine articles. There are also collections of e-books, e-journals available from the CityU Library.)

1.	Rizzo, M. L. (2007). *Statistical computing with R*. Chapman and Hall/CRC.
2.	Givens, G. H., & Hoeting, J. A. (2012). *Computational statistics* (Vol. 710). John Wiley & Sons.

2.2 Additional Readings

(Additional references for students to learn to expand their knowledge about the subject.)