

SDSC3005: COMPUTATIONAL STATISTICS

Effective Term

Semester A 2022/23

Part I Course Overview

Course Title

Computational Statistics

Subject Code

SDSC - School of Data Science

Course Number

3005

Academic Unit

School of Data Science (DS)

College/School

School of Data Science (DS)

Course Duration

One Semester

Credit Units

3

Level

B1, B2, B3, B4 - Bachelor's Degree

Medium of Instruction

English

Medium of Assessment

English

Prerequisites

SDSC3007 Advanced Statistics

Precursors

Nil

Equivalent Courses

Nil

Exclusive Courses

Nil

Part II Course Details

Abstract

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.

Course Intended Learning Outcomes (CILOs)

CILOs		Weighting (if DEC-A1 app.)		DEC-A2	DEC-A3
1	Understand the application background of statistical computing algorithm and techniques.	10	x		
2	Elaborate the theories behind the algorithms in computational statistics.	20	x	x	
3	Implement various types of statistical computing algorithms in R.	20		x	x
4	Apply the correct algorithm to solve a statistical computing problem.	30	x	x	x
5	Discuss tuning parameters for various statistical computing algorithms.	20	x	x	

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 real-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.

Teaching and Learning Activities (TLAs)

TLAs		Brief Description	CILO No.	Hours/week (if applicable)
1	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.	1, 2, 3, 4, 5	26 hours/semester

2	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.	3, 4, 5	13 hours/semester
---	--------------------	---	---------	-------------------

Assessment Tasks / Activities (ATs)

	ATs	CILO No.	Weighting (%)	Remarks (e.g. Parameter for GenAI use)
1	Test	2, 3, 4	30	
2	Assignments	1, 2, 3, 4, 5	30	

Continuous Assessment (%)

60

Examination (%)

40

Examination Duration (Hours)

2

Assessment Rubrics (AR)**Assessment Task**

Test

Criterion

2-hour test to assess students' understanding of computational statistics methods and algorithms.

Excellent (A+, A, A-)

High

Good (B+, B, B-)

Significant

Fair (C+, C, C-)

Moderate

Marginal (D)

Basic

Failure (F)

Not even reaching marginal levels

Assessment Task

Assignments

Criterion

Students' ability to correctly apply computational statistics methods in R to solve given statistics problems.

Excellent (A+, A, A-)

High

Good (B+, B, B-)

Significant

Fair (C+, C, C-)

Moderate

Marginal (D)

Basic

Failure (F)

Not even reaching marginal levels

Assessment Task

Examination

Criterion

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.

Excellent (A+, A, A-)

High

Good (B+, B, B-)

Significant

Fair (C+, C, C-)

Moderate

Marginal (D)

Basic

Failure (F)

Not even reaching marginal levels

Additional Information for AR

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

Part III Other Information

Keyword Syllabus

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

Reading List

Compulsory Readings

Title	
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.

Additional Readings

Title	
1	Nil