# SDSC8004: ADVANCED STATISTICS

#### **Effective Term**

Semester B 2024/25

# Part I Course Overview

#### **Course Title**

**Advanced Statistics** 

# **Subject Code**

SDSC - Data Science

#### Course Number

8004

# **Academic Unit**

Data Science (DS)

#### College/School

College of Computing (CC)

#### **Course Duration**

One Semester

# **Credit Units**

3

#### Level

R8 - Research Degree

# **Medium of Instruction**

English

# **Medium of Assessment**

English

# **Prerequisites**

Nil

# Precursors

Nil

# **Equivalent Courses**

Nil

#### **Exclusive Courses**

Nil

# Part II Course Details

#### **Abstract**

This course aims to provide students with a solid foundation of statistical concepts, theory, and methods including probability theory, statistical estimation and inference methods, and multivariate statistics. It also aims to provide students

with a rigorous introduction to the theory and implementation of linear regression models. Emphasis will be placed on rigorous mathematical derivations of the fundamentals of statistics but implementation of the statistical methods via computer programming in MATLAB or R will be an important part of the course as well.

### **Course Intended Learning Outcomes (CILOs)**

	CILOs	Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Prove and apply various fundamental results in probability theory.	20	x	X	
2	Implement the techniques of parametric inference such as maximum likelihood estimation and Bayesian inference.	35	x	x	
3	Derive key results in the theory of linear models and linear model selection.	25	x	X	
4	Derive key results in the theory of nonparametric statistical models and methods such as the bootstrap method and Gaussian process regression.	10	x	x	
5	Implement statistical inference methods and modelling methodologies with computer codes.	10	х	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.

# Learning and Teaching Activities (LTAs)

	LTAs	Brief Description	CILO No.	Hours/week (if applicable)
1	Lecture	Students will engage in formal lectures to gain knowledge about the theory and methods of advanced statistics.	1, 2, 3, 4, 5	26 hours/semester
2	Demonstration of computer codes	Students will develop an understanding of the computer codes included in the course materials by following an in-class demonstration and explanation of the codes.	1, 2, 3, 4, 5	12 hours/semester

# Assessment Tasks / Activities (ATs)

	ATs	CILO No.	Weighting (%)	Remarks (e.g. Parameter for GenAI use)
1	Midterm Students will be assessed via the midterm their understanding of concepts theory, and methods learned in class, textbooks, and reading materials.	1, 2, 3, 4	25	
2	Two assignments Students will work individually to derive or prove results in probability and statistical theory, and apply statistical methods to analyse data with the help of software.	1, 2, 3, 4, 5	50	
3	Examination Students will be assessed via the examination their understanding of concepts, theory, and methods learned in class, textbooks, and reading materials.	1, 2, 3, 4		

# Continuous Assessment (%)

75

# **Examination (%)**

25

# **Examination Duration (Hours)**

3

# **Additional Information for ATs**

Examination

Students will be assessed via the examination their understanding of concepts theory, and methods learned in class, textbooks, and reading materials.

# Assessment Rubrics (AR)

# **Assessment Task**

Midterm exam (for students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter)

#### Criterion

Midterm exam to assess students' level of achievement of CILOs 1-4 on material covered before the midterm exam.

# **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 (for students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter)

#### Criterion

Assignments are designed to assess student's level of achievement of CILOs 1-5.

#### 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 (for students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter)

#### Criterion

Examination questions are designed to assess student's level of achievement of CILOs 1-4 on all material covered. Students will need to demonstrate understanding of various elements of statistical theory and methods taught in the course through precise mathematical exposition.

#### **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**

Midterm exam (for students admitted from Semester A 2022/23 to Summer Term 2024)

#### Criterion

Midterm exam to assess students' level of achievement of CILOs 1-4 on material covered before the midterm exam.

#### **Excellent**

(A+, A, A-) High

#### Good

(B+, B) Moderate

#### Marginal

(B-, C+, C) Basic

#### **Failure**

(F) Not even reaching marginal levels

#### Assessment Task

Assignments (for students admitted from Semester A 2022/23 to Summer Term 2024)

#### Criterion

Assignments are designed to assess student's level of achievement of CILOs 1-5.

#### **Excellent**

(A+, A, A-) High

#### Good

(B+, B) Moderate

#### Marginal

(B-, C+, C) Basic

### Failure

(F) Not even reaching marginal levels

#### Assessment Task

Examination (for students admitted from Semester A 2022/23 to Summer Term 2024)

#### Criterion

Examination questions are designed to assess student's level of achievement of CILOs 1-4 on all material covered. Students will need to demonstrate understanding of various elements of statistical theory and methods taught in the course through precise mathematical exposition.

#### **Excellent**

(A+, A, A-) High

#### Good

(B+, B) Moderate

# Marginal

(B-, C+, C) Basic

#### **Failure**

(F) Not even reaching marginal levels

# Part III Other Information

# **Keyword Syllabus**

- Probability theory and distributions (probability space, random variables, expectation, inequalities, and convergence of random variables)
- Parametric statistical inference theory and methods (maximum likelihood estimation, Fisher's scoring, Fisher information, consistency and limiting distribution of maximum likelihood estimators, statistical decision theory, Rao-Blackwell theorem, minimum variance unbiased estimation, Bayesian inference)
- Multivariate statistics (covariance matrix estimation, James-Stein estimator, principle components analysis), linear model theory (least squares, Gauss Markov theorem, ridge regression, leave-one-out cross validation, optimal design of experiments), variable selection methods (Bayesian information criterion, LASSO, LARS).
- Nonparametric statistical models and methods (bootstrap, Gaussian process models, local polynomial regression, kernel methods)

#### **Reading List**

# **Compulsory Readings**

	Title
1	Wasserman, L. (2013). All of statistics: a concise course in statistical inference. Springer Science & Business Media.
2	Keener, R. W. (2011). Theoretical statistics: Topics for a core course. Springer.
3	Resnick, S. I. (2013). A probability path. Springer Science & Business Media.
4	Casella, G., & Berger, R. L. (2002). Statistical inference (Vol. 2). Pacific Grove, CA: Duxbury.
5	Rasmussen, C. E., & Williams, C. K. (2006). Gaussian Process Regression for Machine Learning. The MIT Press

# **Additional Readings**

	Title	
1	Nil	