MA5617: STATISTICAL DATA ANALYSIS

Effective Term
Semester B 2024/25

Part I Course Overview

Course Title
Statistical Data Analysis

Subject Code
MA - Mathematics
Course Number
5617

Academic Unit
Mathematics (MA)

College/School
College of Science (SI)

Course Duration
One Semester

Credit Units
3

Level
P5, P6 - Postgraduate Degree

Medium of Instruction
English

Medium of Assessment
English

Prerequisites
Nil

Precursors
Nil

Equivalent Courses
Nil

Exclusive Courses
Nil

Part II Course Details

Abstract
Statistical data analysis in financial business often involves with using sample data to investigate relationships between financial variables and instruments, with an ultimate goal of creating a statistical model for future prediction. This course

offers an introduction to a wide spectrum of statistical modelling techniques, ranging from linear regression, ANOVA, model selection, logistic regression, to nonlinear and nonparametric models.

Course Intended Learning Outcomes (CILOs)

	CILOs	Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Explain the assumptions and context for linear regression, and use it to estimate and predict likely values.	25	x	x	
2	Be able to create appropriate regression models based on data description.	25	x	x	
3	Explain how categorical predictors can be included into a regression model and the different ways of coding the categorical predictors.	25	x	x	
4	Identify strategies to transform data in order to deal with problems identified in the regression model, perform model assessment typically encountered in regression contexts.	25	x	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.

Learning and Teaching Activities (LTAs)

	LTAs	Brief Description	CILO No.	Hours/week (if applicable)
1	Teaching	Students will gain knowledge and participate in discussions on regression techniques	1, 2, 3, 4	3 hours/week

Assessment Tasks / Activities (ATs)

	ATs	CILO No.	Weighting (%)	Remarks (e.g. Parameter for GenAI use)
1	Hand-in assignments	1, 2, 3, 4	20	
2	Project	1, 2, 3, 4	20	

Continuous Assessment (%)

40

Examination (%)

60

Examination Duration (Hours)

2

Additional Information for ATs

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

Assessment Rubrics (AR)

Assessment Task

1. Hand-in assignments (for students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter)

Criterion

1.1 Comprehensive description of the mathematical procedure of estimation as well as inferences problems in linear regression.

Excellent

(A+, A, A-) Consistently demonstrates a thorough understanding of estimation and inferences in linear regression and has strong ability to solve complex problems using R

Good

(B+, B, B-) Adequately demonstrates an understanding of estimation and inferences in linear regression and has ability to solve complex problems using R

Fair

(C+, C, C-) Demonstrates some understanding of estimation and inferences in linear regression and has some ability to solve simple problems using R

Marginal

(D) Demonstrates limited understanding of estimation and inferences in linear regression and has limited ability to solve simple problems using R

Failure

(F) Demonstrates little understanding of estimation and inferences in linear regression and is unable to solve relevant problems

---

Assessment Task

1. Hand-in assignments (for students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter)

Criterion

1.2 Ability to apply appropriate statistical tests to test the stated hypotheses.

Excellent

(A+, A, A-) Consistently demonstrates a thorough understanding of hypothesis testing and has strong ability to solve complex problems using R

Good

(B+, B, B-) Adequately demonstrates an understanding of hypothesis testing and has ability to solve complex problems using R

Fair

(C+, C, C-) Demonstrates some understanding of hypothesis testing and has some ability to solve simple problems using R

Marginal

(D) Demonstrates limited understanding of hypothesis testing and has limited ability to solve simple problems using R

Failure

(F) Demonstrates little understanding of hypothesis testing and is unable to solve relevant problems

---

Assessment Task

1. Hand-in assignments (for students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter)

Criterion

1.3 Ability to interpret the results of hypothesis testing, including p-values and confidence intervals.

Excellent

(A+, A, A-) Consistently demonstrates a thorough understanding of p-values and confidence intervals and has strong ability to solve complex problems using R

Good

(B+, B, B-) Adequately demonstrates an understanding of p-values and confidence intervals and has ability to solve complex problems using R

Fair

(C+, C, C-) Demonstrates some understanding of p-values and confidence intervals and has some ability to solve simple problems using R

Marginal

(D) Demonstrates limited understanding of p-values and confidence intervals and has limited ability to solve simple problems using R

Failure

(F) Demonstrates little understanding of p-values and confidence intervals and is unable to solve relevant problems

---

Assessment Task

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

Criterion

2.1 Ability to gather an appropriate dataset and conduct necessary data preprocessing.

Excellent

(A+, A, A-) Demonstrates a comprehensive understanding of dataset and data preprocessing and strong ability in applying R to solve complex problems

Good

(B+, B, B-) Adequately demonstrates an understanding of dataset and data preprocessing and ability in applying R to solve relevant problems

Fair

(C+, C, C-) Demonstrates some understanding of dataset and data preprocessing and little ability in applying R to solve simple problems

Marginal

(D) Demonstrates limited understanding of dataset and data preprocessing but cannot apply R to solve simple problems

Failure

(F) Inappropriately or unable to apply data analysis using R to solve problems

---

Assessment Task

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

Criterion

2.2 Ability to apply the statistical techniques and concepts covered in the course to build a suitable model for the dataset.

Excellent

(A+, A, A-) Demonstrates a comprehensive understanding of statistical techniques and modelling and strong ability in applying R to solve complex problems

Good

(B+, B, B-) Adequately demonstrates an understanding of statistical techniques and modelling and ability in applying R to solve relevant problems

Fair

(C+, C, C-) Demonstrates some understanding of statistical techniques and modelling and little ability in applying R to solve simple problems

Marginal

(D) Demonstrates limited understanding of statistical techniques and modelling but cannot apply R to solve simple problems

Failure

(F) Inappropriately or unable to apply data analysis using R to solve problems

---

Assessment Task

2. Project (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion

2.3 Ability to interpret and discuss the results of the fitted models and comment on the limitations of the chosen model.

Excellent

(A+, A, A-) Demonstrates a comprehensive understanding of modelling and result evaluation and strong ability in applying R to solve complex problems

Good

(B+, B, B-) Adequately demonstrates an understanding of modelling and result evaluation and ability in applying R to solve relevant problems

Fair

(C+, C, C-) Demonstrates some understanding of modelling and result evaluation and little ability in applying R to solve simple problems

Marginal

(D) Demonstrates limited understanding of modelling and result evaluation but cannot apply R to solve simple problems

Failure

(F) Inappropriately or unable to apply data analysis using R to solve problems

---

Assessment Task

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

Criterion

3.1 Explain the fundamental concepts, principles, and their application scenarios.

Excellent

(A+, A, A-) Consistently demonstrates a thorough understanding of the concepts, principles, and applications of data analysis and has strong ability to solve complex problems

Good

(B+, B, B-) Adequately demonstrates an understanding of the concepts, principles, and applications of data analysis and has ability to solve complex problems

Fair

(C+, C, C-) Demonstrates some understanding of the concepts, principles, and applications of data analysis and has some ability to solve simple problems

Marginal

(D) Demonstrates limited understanding of the concepts, principles, and applications of data analysis and has limited ability to solve simple problems

Failure

(F) Demonstrates little understanding of the concepts, principles, and applications of data analysis and is unable to solve relevant problems

---

Assessment Task

3. Examinations (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion

3.2 Ability to solve modelling related problems with appropriate methods.

Excellent

(A+, A, A-) Consistently demonstrates a thorough understanding of statistical models and methods and has strong ability to solve complex problems

Good

(B+, B, B-) Adequately demonstrates an understanding of statistical models and methods and has ability to solve complex problems

Fair

(C+, C, C-) Demonstrates some understanding of statistical models and methods and has some ability to solve simple problems

Marginal

(D) Demonstrates limited understanding of statistical models and methods and has limited ability to solve simple problems

Failure

(F) Demonstrates little understanding of statistical models and methods and is unable to solve relevant problems

---

Assessment Task

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

Criterion

3.3 Ability to evaluate the model based on the computed solution and use suitable visualizations to effectively present the results.

Excellent

(A+, A, A-) Consistently demonstrates a thorough understanding of data analysis and has strong ability to solve complex problems

Good

(B+, B, B-) Adequately demonstrates an understanding of data analysis and has ability to solve complex problems

Fair

(C+, C, C-) Demonstrates some understanding of data analysis and has some ability to solve simple problems

Marginal

(D) Demonstrates limited understanding of data analysis and has limited ability to solve simple problems

Failure

(F) Demonstrates little understanding of data analysis and is unable to solve relevant problems

---

Assessment Task

1. Hand-in assignments (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion

1.1 Comprehensive description of the mathematical procedure of estimation as well as inferences problems in linear regression.

Excellent

(A+, A, A-) Consistently demonstrates a thorough understanding of estimation and inferences in linear regression and has strong ability to solve complex problems using R

Good

(B+, B) Adequately demonstrates an understanding of estimation and inferences in linear regression and has ability to solve complex problems using R

Marginal

(B-, C+, C) Demonstrates some understanding of estimation and inferences in linear regression and has some ability to solve simple problems using R

Failure

(F) Demonstrates little understanding of estimation and inferences in linear regression and is unable to solve relevant problems

---

Assessment Task

1. Hand-in assignments (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion

1.2 Ability to apply appropriate statistical tests to test the stated hypotheses.

Excellent

(A+, A, A-) Consistently demonstrates a thorough understanding of hypothesis testing and has strong ability to solve complex problems using R

Good

(B+, B) Adequately demonstrates an understanding of hypothesis testing and has ability to solve complex problems using R

Marginal

(B-, C+, C) Demonstrates some understanding of hypothesis testing and has some ability to solve simple problems using R

Failure

(F) Demonstrates little understanding of hypothesis testing and is unable to solve relevant problems

---

Assessment Task

1. Hand-in assignments

Criterion

1.3 Ability to interpret the results of hypothesis testing, including p-values and confidence intervals.

Excellent

(A+, A, A-) Consistently demonstrates a thorough understanding of p-values and confidence intervals and has strong ability to solve complex problems using R

Good

(B+, B) Adequately demonstrates an understanding of p-values and confidence intervals and has ability to solve complex problems using R

Marginal

(B-, C+, C) Demonstrates some understanding of p-values and confidence intervals and has some ability to solve simple problems using R

Failure

(F) Demonstrates little understanding of p-values and confidence intervals and is unable to solve relevant problems

---

Assessment Task

2. Project (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion

2.1 Ability to gather an appropriate dataset and conduct necessary data preprocessing.

Excellent

(A+, A, A-) Demonstrates a comprehensive understanding of dataset and data preprocessing and strong ability in applying R to solve complex problems

Good

(B+, B) Adequately demonstrates an understanding of dataset and data preprocessing and ability in applying R to solve relevant problems

Marginal

(B-, C+, C) Demonstrates some understanding of dataset and data preprocessing and little ability in applying R to solve simple problems

Failure

(F) Inappropriately or unable to apply data analysis using R to solve problems

---

Assessment Task

2. Project (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion

2.2 Ability to apply the statistical techniques and concepts covered in the course to build a suitable model for the dataset.

Excellent

(A+, A, A-) Demonstrates a comprehensive understanding of statistical techniques and modelling and strong ability in applying R to solve complex problems

Good

(B+, B) Adequately demonstrates an understanding of statistical techniques and modelling and ability in applying R to solve relevant problems

Marginal

(B-, C+, C) Demonstrates some understanding of statistical techniques and modelling and little ability in applying R to solve simple problems

Failure

(F) Inappropriately or unable to apply data analysis using R to solve problems

---

Assessment Task

2. Project (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion

2.3 Ability to interpret and discuss the results of the fitted models and comment on the limitations of the chosen model.

Excellent

(A+, A, A-) Demonstrates a comprehensive understanding of modelling and result evaluation and strong ability in applying R to solve complex problems

Good

(B+, B) Adequately demonstrates an understanding of modelling and result evaluation and ability in applying R to solve relevant problems

Marginal

(B-, C+, C) Demonstrates some understanding of modelling and result evaluation and little ability in applying R to solve simple problems

Failure

(F) Inappropriately or unable to apply data analysis using R to solve problems

---

Assessment Task

3. Examinations (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion

3.1 Explain the fundamental concepts, principles, and their application scenarios.

Excellent

(A+, A, A-) Consistently demonstrates a thorough understanding of the concepts, principles, and applications of data analysis and has strong ability to solve complex problems

Good

(B+, B) Adequately demonstrates an understanding of the concepts, principles, and applications of data analysis and has ability to solve complex problems

Marginal

(B-, C+, C) Demonstrates some understanding of the concepts, principles, and applications of data analysis and has some ability to solve simple problems

Failure

(F) Demonstrates little understanding of the concepts, principles, and applications of data analysis and is unable to solve relevant problems

---

Assessment Task

3. Examinations (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion

3.2 Ability to solve modelling related problems with appropriate methods.

Excellent

(A+, A, A-) Consistently demonstrates a thorough understanding of statistical models and methods and has strong ability to solve complex problems

Good

(B+, B) Adequately demonstrates an understanding of statistical models and methods and has ability to solve complex problems

Marginal

(B-, C+, C) Demonstrates some understanding of statistical models and methods and has some ability to solve simple problems

Failure

(F) Demonstrates little understanding of statistical models and methods and is unable to solve relevant problems

---

Assessment Task

3. Examinations (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion

3.3 Ability to evaluate the model based on the computed solution and use suitable visualizations to effectively present the results.

Excellent

(A+, A, A-) Consistently demonstrates a thorough understanding of data analysis and has strong ability to solve complex problems

Good

(B+, B) Adequately demonstrates an understanding of data analysis and has ability to solve complex problems

Marginal

(B-, C+, C) Demonstrates some understanding of data analysis and has some ability to solve simple problems

Failure
(F) Demonstrates little understanding of data analysis and is unable to solve relevant problems

---

Part III Other Information

Keyword Syllabus
linear regression; ordinary least squares; ANOVA; model selection; logistic regression; nonlinear regression; smoothing

Reading List

Compulsory Readings

	Title
1	Applied Linear Statistical Models by Kutner, Nachtsheim, Neter, and Li, McGraw-Hill Irwin, 2005

Additional Readings

	Title
1	The Statistical Sleuth by Ramsey and Schafer, 3rd Edition, Cengage Learning