Macao Polytechnic University

Faculty of Applied Sciences

Master of Science in Big Data and Internet of Things

Module Outline

Academic Year   2022 / 2023   Semester 2

Learning Module	Advanced topics in probability and statistics		Class Code	COMP6105
Pre-requisite(s)	Nil			
Medium of Instruction	English		Credit	3
Lecture Hours	45 hrs	Lab/Practice Hours	0 hrs	Total Hours	45 hrs
Instructor	Dennis Wong Chi Him		E-mail	cwong@mpu.edu.mo
Office	Rm. N46B, Wui Chi Building, Main Campus		Telephone	85996875

Description

This module will cover core concepts of probability theory and statistics with applications. Specific topics will include random variables and distributions, quantitative research methods (correlation and regression), and modern techniques of optimization and machine learning (clustering and prediction).

Learning Outcomes

After completing the learning module, students will be able to:

1. Explain clearly concepts from advanced probability and stochastic processes; (M3)

2. Develop techniques needed to accurately calculate probabilities; (M2, M3)

3. Apply probability theory to solve real-world problems; (M1, M3)

4. Analyze and interpret statistical data using appropriate probability distributions. (M1, M2, M3)

Content

1. Probability Distributions   (9 hours)

    1.1   Bernoulli distribution

    1.2   Binomial distribution

    1.3   Poisson distribution

    1.4   Gaussian distribution

    1.5   Chi-squared distribution

    1.6   Independence, joint, and conditional distributions

    1.7   Expectation and variance

2. Decision Theory   (6 hours)

    2.1   Minimizing expected loss

    2.2   Loss functions

    2.3   Classification

    2.4   Prediction

    2.5   Bayes rules and Bayesian inference

3. Probability Inequalities   (3 hours)

    3.1   Chebyshev's inequality

    3.2   Chernoff inequality

    3.3   Law of large numbers

4. Moment Generating Functions   (3 hours)

    4.1   Moment generating functions

    4.2   Central limit theorem

    4.3   Relationship to statistical inference

5. Point and Interval Estimation   (3 hours)

    5.1   Method-of-moments

    5.2   Maximum likelihood

    5.3   Consistency and asymptotic normality

6. Hypothesis Testing   (6 hours)

    6.1   Type I and type II error

    6.2   Familywise error rate

    6.3   False discovery rate

    6.4   Benjamini-Hochberg procedure

7. Linear Regression   (9 hours)

    7.1   Least-squares

    7.2   Bias-variance decomposition

    7.3   Bayesian linear regression

    7.4   Fisher's linear discriminant

8. Non-linear Regression and Classification   (6 hours)

    8.1   Kernel density estimators

    8.2   Nearest-neighbor classifiers

Teaching Method

Lectures and tutorials.

Attendance

Attendance requirements are governed by the "Academic Regulations Governing Master's Degree Programmes" of Macao Polytechnic University. Students who do not meet the attendance requirements for the module will not be permitted to sit the final examination and shall be awarded an 'F' grade.

Assessment

This learning module is graded on a 100 point scale, with 100 being the highest possible score and 50 the pass score.

	Item	Description	AHEP4 LO	Percentage
1.	Assignments	Knowledge assessment	M1, M2, M3	30%
2.	Test	Knowledge assessment	M1, M2, M3	30%
3.	Examination	3-hour written examination	M1, M2, M3	40%
	Total Percentage:			100%

Students with an overall score of less than 35 in the coursework will fail the module even if the overall score for the module is 50 or above.

Students with a score of less than 35 in the final examination will fail the module even if the overall score for the module is 50 or above.

Teaching Material(s)

Textbook(s)

Wasserman, Larry. "All of Statistics: A Concise Course in Statistical Inference", 1st edition, Springer New York, 2004.

Reference

Reference book(s)