

**MATH 3424**  
**Regression Analysis**  
**2024-25 Spring**  
<https://canvas.ust.hk/courses/62302>

| LECTURE                   |                                                                                |
|---------------------------|--------------------------------------------------------------------------------|
| <b>Time</b>               | Tue & Thu 15:00 - 16:20                                                        |
| <b>Venue</b>              | Room 2464                                                                      |
| <b>Instructor</b>         | <b>Prof. Dong XIA</b>                                                          |
| <b>E-mail</b>             | madxia@ust.hk ( <i>send an email to make an appointment</i> )                  |
| <b>Office</b>             | Room 3431, Department of Mathematics                                           |
| TUTORIAL -T1A             |                                                                                |
| <b>Time &amp; Venue</b>   | Wed 15:00-15:50, Room 1410                                                     |
| <b>Teaching Assistant</b> | <b>Mr. Su YAN</b>                                                              |
| <b>E-mail</b>             | syanal@connect.ust.hk ( <i>send an email to make an appointment</i> )          |
| <b>Programming</b>        | (TA will give several R programming tutorial sessions in the first few weeks.) |
| TUTORIAL -T1B             |                                                                                |
| <b>Time &amp; Venue</b>   | Tue 12:00-12:50, Room 2302                                                     |
| <b>Teaching Assistant</b> | <b>Ms. Congyuan DUAN</b>                                                       |
| <b>E-mail</b>             | cduanac@connect.ust.hk ( <i>send an email to make an appointment</i> )         |
| <b>Programming</b>        | (TA will give several R programming tutorial sessions in the first few weeks.) |

COURSE DESCRIPTION

**Course outline:** Estimation and hypothesis testing in linear regression, residual analysis, multicollinearity, indicator variables, variable selection, non-linear regression, logistic regression.

**Credits:** 3

**Prerequisites:** Probability (MATH 2421) and Statistical Inference (MATH 3423).

INTENDED LEARNING OUTCOMES (ILOs)

Upon completion of this course, students are expected to:

- (1) Get familiar with simple and multiple linear regression, the involved statistical inference, prediction, the measure of goodness-of-fit, and applications on real data examples;
- (2) Understand the diagnostics of regression, residuals, checking linearity, leverage, influence, dealing with outliers and applications on real data examples;
- (3) Understand how to treat categorical variables as predictors, the transformation of variables, treating heteroscedastic errors and applications on real data examples;
- (4) Understand variable selection, the procedures and applications;
- (5) Understand logistic regression for classification problem, the quality of fit, determination of important variables and applications on real data examples.

ASSESSMENT AND GRADING

**Homework:** There will be 5 Assignments. Students should submit each homework in form of a **clearly written and scanned** or a **LaTeX-typed** PDF on the **Canvas** system before the deadline. Late homework will NOT be accepted. Team work is NOT allowed!

**Examinations:** There will be a 85-minute midterm exam during lecture time(exact date to be confirmed), and a 3-hour final exam arranged by ARO.

Make-up midterm policy:

- Under any circumstance, students who are unable to attend the midterm exam will NOT be offered a make-up midterm.

Final exam policy:

- The course will follow the make-up exam policy set by ARO for the final exam. Approval from the instructor, the dean, and ARO will be needed for applying for a make-up final exam.

### Grading Scheme:

This course will be assessed using **criterion-referencing**, and grades will **not** be assigned using a curve. Your course total will be calculated by taking the **maximum of two schemes** in order to encourage students to work harder in the final exam in case the midterm result is not desirable or some student missed the midterm exam:

|                     | Scheme A | Scheme B | Address ILOs |
|---------------------|----------|----------|--------------|
| Homework            | 15%      | 15%      | 1, 2, 3,4,5  |
| Midterm             | 30%      | 0%       | 1, 2, 3,4,5  |
| Final               | 55%      | 85%      | 1, 2, 3,4,5  |
| <b>Course Total</b> | 100%     | 100%     |              |

**Letter Grades:** Students should aim at getting a course total of 85% or above for A-/A/A+, and about 60% or above for B-/B/B+.

### Grade Descriptors:

| Grades | Short Description        | Elaboration on subject grading description                                                                                                                                                                                                              |
|--------|--------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| A      | Excellent Performance    | The student has mastered almost all concepts and techniques of regression analysis taught in the course, has excellent understanding of the deepest content of the subject, and acquired workable knowledge for further studies of advanced statistics. |
| B      | Good Performance         | The student has mastered the application of the statistical methods taught in this course, yet the understanding of some challenging concepts may not be deep enough for further studies on related advanced subjects.                                  |
| C      | Satisfactory Performance | The student meets the minimum expectation of the instructor, has acquired some basic knowledge of the taught statistical methods, yet some concepts were not clearly understood.                                                                        |
| D      | Marginal Pass            | The student is only able to recall some fragments of topics and is able to apply some of the easiest statistical methods                                                                                                                                |
| F      | Fail                     | The student does not have sufficient understanding of even some fragments of topics, and is not even able to apply some of the easiest statistical methods.                                                                                             |

### Course AI Policy

The use of ChatGPT or other generative AI is allowed for Assignments ONLY, but needs to be declared in your submissions.

Please be warned that at the current stage of development of AI, the response to problems in advanced statistics courses is not always reliable. Students should be critical of the response generated by AI and do not blindly copy the generated responses to your homework.

### TEXTBOOK AND LEARNING RESOURCES

**Lecture notes and slides** (on canvas)

**Textbook:** Chatterjee and Hadi, Regression Analysis by Example, Fifth Edition (*not required*).  
David Dalpiaz, Applied Statistics with R ([available on GitHub](#), [google it](#)).

## ACADEMIC INTEGRITY

Students are expected to adhere to the university's academic integrity policy. Students are expected to uphold HKUST's Academic Honor Code and to maintain the highest standards of academic integrity. The University has zero tolerance of academic misconduct.