

The Chinese University of Hong Kong
Faculty of Science
Science Academy for Young Talent

Summer Courses 2023
Course Outline

CUSA1046 Patterns of Real Data and Simulated Data
真數據和模擬數據的規律

Introduction: Statistics turns real data into science. It can also define rules to generate simulated "fake" data to understand nature. This course discusses both approaches of statistical analysis. Four themes are covered: (1) privacy protection, (2) stock price prediction, (3) trend inference of global temperature, and (4) cancer image data analysis. The problem in each theme is resolved by one advanced statistical methodology. The tools include randomized estimator, historical simulation, nonparametric trend estimation, and Kolmogorov–Smirnov test for distribution comparison. All topics are supplemented with computing laboratory sessions.

統計學將真實數據轉化為科學。它還可以定義規則來生成模擬的「假數據」以了解自然。本課程討論以上兩種統計分析的方法，並涵蓋四個大主題：(一) 隱私保護、(二) 股票價格預測、(三) 全球氣溫趨勢和(四) 癌症圖像數據分析。每個主題的問題都通過一種先進的統計方法來解決。這些工具包括隨機估計器、歷史模擬、非參數趨勢估計和用於分佈比較的 Kolmogorov – Smirnov 檢驗。所有主題都輔以計算實驗室課程。

Medium of Instruction: Cantonese supplemented with English
粵語輔以英語

Organising Unit: Department of Statistics, Faculty of Science, CUHK

Teachers:



Prof. CHAN, Kin Wai (陳健威教授)

Assistant Professor

Department of Statistics, CUHK

Room 115, Lady Shaw Building, CUHK

E-mail: kinwaichan@cuhk.edu.hk

Course Content:

24 July 2023 (Monday) 9:30 am – 12:30 pm 2:00 pm – 5:00 pm	<p><u>Theme 1: Privacy protection 隱私保護</u></p> <ul style="list-style-type: none"> Theory of estimation. 估算理論。 Qualities of estimators. 估計量的素質。 Construction of privacy-protected procedures. 構建隱私保護程序。 Laboratory session: statistical programming and simulation experiments. 實驗室課程：統計電腦程序和模擬實驗。
25 July 2023 (Tuesday) 9:30 am – 12:30 pm 2:00 pm – 5:00 pm	<p><u>Theme 2: Stock price prediction 股價預測</u></p> <ul style="list-style-type: none"> Theory of resampling. 重採樣理論。 Sampling and resampling. 採樣和重採樣程序。 Historical simulation. 歷史模擬法。 Laboratory session: stock price prediction and stock trading experiment. 實驗室課程：股票價格預測和股票交易實驗。
26 July 2023 (Wednesday) 9:30 am – 12:30 pm 2:00 pm – 5:00 pm	<p><u>Theme 3: Trend of global temperature 全球氣溫趨勢</u></p> <ul style="list-style-type: none"> Theory of regression. 回歸理論。 Best fitted line and parametric method. 最佳迴歸線和參數方法。 Best fitted curve and nonparametric method. 最佳迴歸曲線和非參數方法。 Laboratory session: global temperature analysis and discussion. 實驗室課程：全球溫度分析和討論。
27 July 2023 (Thursday) 9:30 am – 12:30 pm 2:00 pm – 5:00 pm	<p><u>Theme 4: Cancer image data analysis 癌症影像數據分析</u></p> <ul style="list-style-type: none"> Theory of testing. 統計測試理論。 Hypotheses and test statistics. 假設和檢驗統計量。 Empirical distribution function and Kolmogorov–Smirnov test. 經驗分佈函數和 Kolmogorov – Smirnov 檢驗。 Laboratory session: cancer image data analysis and discussion. 實驗室課程：癌症圖像數據分析和討論。
28 July 2023* (Friday) 9:30 am – 12:30 pm 2:00 pm – 5:00 pm	Make-up Class

Duration	4 whole day sessions (total 24 contact hours)
Date	24 – 27 July 2023 28 July 2023* (make up class)
Time	9:30 am – 12:30 pm & 2:00 pm – 5:00 pm
Teaching Mode	Face to Face (The Chinese University of Hong Kong)
Enrollment	30
Expected Applicants	Students who are promoting to or studying S4 – S6
Tuition Fee	HKD 3,540.00
Credit	1.5 Academy Unit Certificate of completion will be awarded to students who pass the assessment (if applicable) and attain at least 75% attendance.
Grading Methods	Letter grades range from A to F.

* This date is reserved for make-up classes in case there is any cancellation of classes due to unexpected circumstances.