國 立 成 功 大 學 模 組 化 課 程

開課學年度/學期：114 學年度第 1 學期　　領域：科際整合

統計與機器學習的數學與哲學基礎
The Mathematical and Philosophical Foundations of Statistics and Machine Learning

教師	任職單位	畢業學校
高翰迪 ika@ucdavis.edu	Philosophy Department, University of California, Davis	卡內基美隆大學

類別	學分數	開課人數	其他注意事項
科際整合	1.5	20	

先修課程或先備能力

無

課程難易度

□難■中偏難 □中偏易 □易

建議修課學生背景

理學院、工學院、生科院、電資學院、醫學院

教學方法

講授 50%，實作 50%

評量方式

問題考試 40 %：禮拜五上午，在課堂考考試
作業 40 %：週一到週四會有回家作業，隔天一開始上課的時候繳交
報告 20 %：實作的過程中，把回家作業的內容報告出來，跟大家分享

學習規範

無

課程概述

尋找奠定科學方法論的基礎，特別關注在如何從數據推論到假說，是一個悠久的哲學傳統。其核心思想是，在評估任何推論方法時，都應考慮其是否在「大多數」情況下引導我們得出真理或者是正確回答我們問的問題。然而，這引發了一個深刻的議題：我們如何能夠嚴格地定義「大多數」，並將這一概念應用於科學方法論的科學研究？這個問題是西方哲學，特別是科學哲學中最深刻的問題之一，並且哲學家、統計學家和機器學習理論家都對其做出了重要貢獻。

在本課程中，我們將探討三種嚴格定義「大多數」的方法：透過集合論中的基數概念 (cardinality)、數理統計中的機率概念 (probability)，以及拓撲學中的貝爾範疇（Baire category）。然後，我們將嘗試運用這些數學框架來分析兩個基本問題：一是科學哲學中的悲觀歸納問題 (pessimistic induction)，二是機器學習領域中如何從非實驗數據推斷因果關係的問題，這遠遠超越了統計學中僅關注相關性的做法。

這門跨學科研究位於數學、科學哲學、統計學與機器學習基礎理論的交匯點，提供了一種獨特的視角來探討西方哲學中最深刻的問題之一。由於本課程的跨學科性質，不同背景的學生將有不同的作業、考試和評分標準。本課程只有一個先修要求：對於理科學生，需熟悉 ε-δ 極限定義；對於哲學學生，需熟悉一階邏輯。

國 立 成 功 大 學 模 組 化 課 程

開課學年度/學期：114 學年度第 1 學期　　領域：科際整合

關鍵字：科學推論、機率、拓樸、統計、機械學習、因果關係

課程概述(英文)

There is a rich philosophical tradition that seeks to ground scientific methodology, particularly regarding the inference from data to hypotheses. At its core is the idea that any method of inference should be evaluated based on its ability to lead us to the truth—or the desired learning target—in "most" cases, if not all. But this raises a profound question: how can we rigorously define "most" and apply this concept to the scientific study of scientific methodology? This question is among the deepest ever explored in Western philosophy, especially in the philosophy of science, with significant contributions from philosophers, statisticians, and machine learning theorists.

In this course, we will explore three rigorous approaches to defining "most": through cardinality, as in set theory; through probabilities, as in mathematical statistics; and through Baire categories, as in topology. We will then try out these mathematical frameworks by applying them to two foundational problems: the challenge of pessimistic induction in the philosophy of science, and the problem of inferring causal relationships in machine learning from non-experimental data, going way beyond mere correlations as usually pursued in statistics.

This interdisciplinary inquiry sits at the intersection of mathematics, philosophy of science, and the foundations of statistics and machine learning, offering a unique perspective on one of philosophy's most enduring questions. Due to the interdisciplinary nature of this course, students with different backgrounds will be given different homework problems, exams, and grading criteria. There is only one prerequisite: familiarity with the epsilon-delta definition of limit (for science students) or first-order logic (for philosophy students).

Keywords：Scientific Inference, Probability, Topology, Statistics, Machine Learning, Causality

課程進度

日期	時間	進 度 說 明
2025/8/11(一)	Morning **9:00-12:00**	**Lecture: A Philosophical Introduction** **- Hume's problem of induction** **- four traditions in the philosophy of scientific inference** **- the tradition of our focus: convergentism**
	12:00-13:00	**Lunch time**
	Afternoon **13:00-15:30**	**Lecture + Discussion Session: A Crash Course on Probability Theory** **- the concept of limit and convergence** **- Bernoulli's weak law of large numbers** **- Hoeffding's inequality** **- (the strong law of large numbers, optional)** **- doing some exercises together** **- goal: to provide the mathematical tools to be used in next morning**

國 立 成 功 大 學 模 組 化 課 程

開課學年度/學期：114 學年度第 1 學期 **領域：科際整合**

2025/8/12(二)	Morning **9:00-12:00**	**Lecture: The Convergentist Foundations of Hypothesis Testing** **- testing deterministic hypotheses (as in formal learning theory)** **- testing statistical hypotheses (as in classical statistics)** **- applications to justification of Ockham's razor in hypothesis testing**
	12:00-13:00	**Lunch time**
	Afternoon **13:00-15:30**	**Discussion Session** **- doing some exercises together, possibly by dividing students with different backgrounds into different groups** **- discussing some philosophical problems underlying the materials covered in the morning**
2025/8/13(三)	Morning **9:00-12:00**	**Lecture: A Crash Course in Set Theory and Topology** **- basic set theory** **- Cantor's diagonal proof that the set of real numbers has a larger cardinality than the set of natural (or rational) numbers** **- the concepts of metric spaces and topological spaces** **- Baire's definition of "big" subsets of a topological space, and the Baire category theorem** **- elementary examples, using the Baire space, the real line, and finite dimensional Euclidean spaces** **- goal: to provide the mathematical tools to be used on the next day**
	12:00-13:00	**Lunch time**
	Afternoon **13:00-15:30**	**Discussion Session** **- again, doing some exercises together** **- discussing why and how set theory and topology may be related to philosophical issues, and anticipate their applications to next morning's materials**
2025/8/14(四)	Morning **9:00-12:00**	**Lecture: The Convergentist Foundations of Causal Inference in Machine Learning** **- walking through the details of a simplest nontrivial problem of choosing a causal model in machine learning, and explaining why both probability and topology need to be used here** **- walking through a surprisingly similar example in the history of science: how exactly Perrin's evidence from the Brownian motion supported the existence of atoms and won him the 1926 Novel Prize in Physics**
	12:00-13:00	**Lunch time**
	Afternoon **13:00-15:30**	**Discussion Session** **- again, doing some exercises together** **- taking some philosophical or mathematical questions from students**

國 立 成 功 大 學 模 組 化 課 程

開課學年度/學期：114 學年度第 1 學期 **領域：科際整合**

2025/8/15(五)	Morning **9:00-12:00**	**Final Exam** **- scheduled for Friday morning, rather than afternoon, because it seems unreasonable to let the final exam cover the materials in the Friday lecture**
	12:00-13:00	**Lunch time**
	Afternoon **13:00-15:30**	**Lecture: The Convergentist Foundations of Scientific Realism** **- addressing Laudan's challenge to the realist philosophy of science: pessimistic induction** **- sketching an unexpected counterpart of Laudan's challenge in machine learning: testing conditional independence can be very hard (which is one of my on-going research projects)** **- wrapping up: what if you prefer not to be a realist in the philosophy of science but instead, say, an instrumentalist?**

課程學習目標

1. **傳統科學哲學議題的掌握**
2. **數學常用的基本工具，如機率與拓撲**
3. **如何把第 2 項應用在第 1 項**

課程的重要性、跨域性與時代性

從 21 世紀初期，開始有大量的哲學家使用數學工具來分析傳統哲學議題。也有科學家使用他們熟悉的數學工具來跟哲學家進行討論。這個進路已經蓬勃發展，開始為現代的統計學以及機械學習提供數學與哲學基礎。台灣目前在這方面參與的學生仍然很少，值得開始推廣。

其他備註

無

本課程若因天災等不可抗力之因素或中央、地方政府公告停課，授課教師需依情況依建議補課方式調整課程進度與補課；若需使用假日、國定假日補課，則需與所有修課學生達成共識方能用例假日補課。

建議補課方式：

1. 線上授課方式補課；
2. 當預期可能會因天災(颱風、超大豪雨…等)宣佈停課時，建議老師先行調整加快課程進度或預先增加可能天氣預警之前幾次課程時數；
3. 停課後隔天起延後下課，補足停課延誤的進度；若停課超過 1 天，則在開始上課後延後下課補課，或當週星期六、日補課；
4. 更改課程授課方式，例如：DEMO 改以考試、報告、作業取代。