

中央研究院統計科學研究所
2017 年統計研習營 簡章

一、宗旨：提供在學學生學習統計相關知識及參與學術研究機會，培育統計領域優秀人才。

二、研習時間：2017 年 7 月 5 日（三）至 7 月 21 日（五）

三、研習地點：中央研究院

四、申請資格：對統計科學有興趣之國內外大學部在學學生（含應屆畢業生）。

五、申請辦法：[線上申請](#)，即日起至 2017 年 6 月 5 日止

六、申請文件：

1. 在學證明（jpg 或 pdf）：學生證（正反面，含本學期註冊章）或其他可證明文件。
2. 歷年成績單（jpg 或 pdf）：含 105 學年度第一學期成績。
3. 填寫並繳交報名表附檔中的四項問題。

七、課程簡介（各課程講師依中文姓氏筆畫排序）

註：研習期間需自備筆電

入門課程（7 月 5 日~7 月 7 日）

◆機率統計課程入門（1.5 天）

黃文璋 教授（國立高雄大學統計學研究所）

羅夢娜 教授（國立中山大學應用數學系）

◆R 語言基礎與統計分析（1 天）

吳漢銘 副教授（國立臺北大學統計學系）

◆Matlab 軟體基礎與應用（0.5 天）

杜憶萍 研究員（中央研究院統計科學研究所）

洪 弘 教授（國立臺灣大學流行病學與預防醫學研究所）

陳定立 副研究員（中央研究院統計科學研究所）

陳素雲 研究員（中央研究院統計科學研究所）

核心課程（7 月 10 日~7 月 14 日、7 月 17 日~7 月 21 日）

◆統計與機器學習（2.5 天）

杜憶萍 研究員（中央研究院統計科學研究所）

洪 弘 教授（國立臺灣大學流行病學與預防醫學研究所）

陳定立 副研究員（中央研究院統計科學研究所）
蔡宗希 副研究員（中央研究院統計科學研究所）
陳素雲 研究員（中央研究院統計科學研究所）

◆工業統計（2.5 天）

王秀瑛 教授（國立交通大學統計學研究所）
洪英超 教授（國立政治大學統計學系）
彭健育 副研究員（中央研究院統計科學研究所）
曾勝滄 教授（國立清華大學統計學研究所）
樊采虹 教授（國立中央大學統計研究所）

◆應用統計與計量（2.5 天）

王馨徽 副教授（國立清華大學計量財務金融學系）

◆社會網絡（2.5 天）

潘建興 副研究員（中央研究院統計科學研究所）
劉維中 副研究員（中央研究院統計科學研究所）
顏佐榕 助研究員（中央研究院統計科學研究所）

八、錄取結果：

本所將依據申請人所備之書面資料，由所方組成之審核小組進行審查並決定錄取人員。本年度預計招收 20 名學員，6 月底以 e-mail 個別通知審查結果，並於活動網頁公告相關資訊。

九、結業證書及研習費：

研習期間經評審成績優良之學員，將頒發結業證書及研習費。

十、住宿：

學員可申請於研習期間自費住宿（7/4 晚~7/20 晚，不含 7/7、7/8、7/14、7/15 四個晚上），由本所代為安排本院學術活動中心雙人房為原則（1,700 元/晚，平均每人負擔 850 元/晚）。

本院學術活動中心客房 <http://gao.sinica.edu.tw/huo-dong/room.php>

十一、聯絡人：賴姿秀小姐、(02) 2783-5611 分機 128、school@stat.sinica.edu.tw

活動網址：<http://www3.stat.sinica.edu.tw/school2017/>

課程簡介

(各課程講師依中文姓氏筆畫排序)

◆機率統計課程入門 (1.5 天)

[黃文璋 教授](#) (國立高雄大學統計學研究所)

[羅夢娜 教授](#) (國立中山大學應用數學系)

由於生活中充滿著隨機性，因此人們常需對現況做推測，或對未來做預測。這些推論如何產生？擲筊？藉助水晶球？雖可有各種推論方式，但科學的推論，仍得仰賴數據，所謂讓數據說話。利用統計方法，先收集資料，加以整理及分析，然後給出推論。

對於隨機現象，即使採用的方法再好，推論仍不易次次準確，也難以保證預測的誤差必有多小。一般人受數學的薰陶較久，數學思維根深蒂固，因而初學統計時，對很多概念不時感到迷惑。例如，數學中常在證明，從給定前提下，推導出結果；而統計則是由觀測到的結果，來判定前提是否可接受。要知在隨機世界裡，真相難明，一切可說都是假設，就看接受那一個。另外，數學中所得結果往往為斬釘截鐵式的，而統計裡所給，乃允許誤差下的機率式保證。隨機性的特質，使統計在理解層次上，並不易掌握。在實際應用時，若未具備充分的統計素養，便難讓統計的功能充分發揮，甚至誤用，使數據有時說錯話，或有時說得太正確。造成有人不相信統計，有人則過度相信統計。只是數據其實並未說話，是人在幫它說話。說得太不準，或說得太準，都該令人起疑，而重新檢視得到推論的過程。

在本課程中，將闡釋隨機性的概念，並介紹一些基本的統計思維。

◆R 語言基礎與統計分析 (1 天)

[吳漢銘 副教授](#) (國立臺北大學統計學系)

R 是資料科學領域裡最受歡迎的程式語言之一，在全球已超過兩佰萬人使用。R 具有免費，開放源碼及學習資源豐富的優點，而其功能可透過眾多套件增強，在生物資訊、經濟計量、財務管理、人文社會科學研究以及人工智慧等等領域已有成功的應用。本研習課程主題為 R 的基礎(資料輸出輸入、函式撰寫、圖形及視覺化)、假設檢定、統計模型與迴歸分析。希望藉由本課程之訓練，引起學員對 R 的學習興趣，奠定學員以 R 語言進行資料統計分析的基礎。

◆Matlab 軟體基礎與應用 (0.5 天)

[杜憶萍 研究員](#) (中央研究院統計科學研究所)

[洪 弘 副教授](#) (國立臺灣大學流行病學與預防醫學研究所)

[陳定立 副研究員](#) (中央研究院統計科學研究所)

[陳素雲 研究員](#) (中央研究院統計科學研究所)

MATLAB (矩陣實驗室) 是資料分析與科學計算常用的程式語言。我們將介紹 MATLAB 的基本環境與操作，並為統計與機器學習課程所需的應用預作準備。

◆統計與機器學習 (2.5 天)

[杜憶萍 研究員](#) (中央研究院統計科學研究所)

[洪 弘 副教授](#) (國立臺灣大學流行病學與預防醫學研究所)

[陳定立 副研究員](#) (中央研究院統計科學研究所)

[陳素雲 研究員](#) (中央研究院統計科學研究所)

隨著知識與科技的蓬勃發展，各種資料不停地大量的產生，如何有效的分析資料，並回饋於真實數據問題，幫助做預測與決策，是統計科學發展的重點與責任。統計機器學習是一重要的資料分析工具，它結合了統計科學與人工智慧，並且已廣泛應用於多種領域。

本課程將介紹監督式學習、非監督式學習，以及它們的應用。我們將提供資料實例與 Matlab 程式，讓學員們能實際動手操作。

◆工業統計 (2.5 天)

[王秀瑛 教授](#) (國立交通大學統計學研究所)

[洪英超 教授](#) (國立政治大學統計學系)

[彭健育 副研究員](#) (中央研究院統計科學研究所)

[曾勝滄 教授](#) (國立清華大學統計學研究所)

[樊采虹 教授](#) (國立中央大學統計研究所)

工業統計一般區分為實驗設計與分析、品質管制以及可靠度分析三大領域，本研習課程將會簡介各領域的研究課題與實際案例，並介紹實際常用的分析方法，相關的理論基礎，讓學員瞭解統計界常用的工具及背後的觀念。

◆應用統計與計量 (2.5 天)

[王馨徽 副教授](#) (國立清華大學計量財務金融學系)

隨著大數據的普遍運用，統計以及計量經濟學...等工具的發展與創新對於資料分析而言，已經變得愈來愈重要。在這個課程中，我們不僅將介紹一些近幾年來各專業領域在建模過程中所發展出來的統計以及計量經濟學方法，我們也將推廣這些方法在財務金融上、在經濟學上、以及在公共政策...等領域的應用。就授課內容而言，主要分成三個部份: (i) 時間序列計量經濟學 (time series econometrics); (ii) 面板數據分析 (panel data analysis); (iii) 以上這兩個學科在非法毒品流量檢測 (illegal drug flow detection)、銀行和金融危機檢測 (banking and financial crises detection)、股票報酬率的預測 (the prediction of stock returns)、以及探討經濟事件和活動對於金融市場的影響...等方面的應用。而這

些議題也是最近研究的焦點，特別是對決策者以及商業從業人員而言，相對重要。舉例來說，如何量測銀行業以及非銀行業系統性風險對於追蹤即將來臨的危機是非常重要的。除此以外，了解非法藥品的價格的決定機制對於監視非法藥物的流動方向也可能會有幫助。為了培養學員的創造力以及開拓其更寬廣的視野，在這個課程中，我們也會安排實例的探究 (case studies)。這些實例探究的目的，主要是讓學員熟悉財經實證主題的模型建構。最後，我們也會介紹 E-views 或是 Stata 這些統計軟體，以用來做一些簡單的資料分析。

◆社會網絡 (2.5 天)

[潘建興 副研究員](#) (中央研究院統計科學研究所)

[劉維中 副研究員](#) (中央研究院統計科學研究所)

[顏佐榕 助研究員](#) (中央研究院統計科學研究所)

當各位網民在臉書上玩遊戲交朋友時，你們知道臉書在後台分析著你們的交友網，以便將來為你們推薦新朋友嗎？當各位戰友在魔獸裡攻城掠地時，你們知道暴雪在後台分析著你們的表現，以便未來更新時為你們設定新的副本嗎？當各位網購族在奇摩商城裡瘋狂血拼時，你們知道雅虎在後台分析著你們的指標行為，以便下次特賣時為你們推薦喜歡的新品嗎？以上所有的分析都離不開當今最熱門的主題之一：社會網絡分析。

本研習將會以實際資料向學員完整地介紹這方面的知識，內容包括介紹網絡的基本定義、構造和特性，如何從點與點的關係量化點的位置以及其重要性，如何把點分群來簡化網絡和不同網絡生成的機制。藉由一些經典例子介紹社會網絡分析的用處和特性，例如人際間同質性現象、策略聯盟、資訊的擴散。接著引進一些基本的網絡概念、模型和統計方法，例如網絡結構、中心性、群聚係數、網絡基序和最小世界模型等。

本研習除了導師們課堂知識的分享外，學員們會以有獎比賽的方式來了解社會網絡形成和變化，務求學員在這 2.5 天的課程中，能夠以最輕鬆的互動來了解如何分析這個當今人類不可或缺的網絡世界。