

一、資料與統計

Chapter 1 Data and Statistics

追尋那道光 李明聰

目錄

一、資料與統計	1
1.1 統計學在管理層面的應用.....	3
1.2 資料	4
1.2.1 相關名詞	4
1.2.2 評量尺度數值類型.....	6
1.2.2.1 名目尺度.....	6
1.2.2.2 順序尺度.....	7
1.2.2.3 等距尺度.....	7
1.2.2.4 比例尺度.....	8
1.2.2.5 問卷設計中數值類型敘述範例.....	8
1.2.3 質化和量化資料.....	10
1.2.4 橫斷面和時間序列資料.....	11
1.3 資料來源	12
1.4 敘述性統計	13
1.5 推論性統計	13
1.6 應用統計學	14
討論議題	16
重點整理	17



學習目標

知識(認知)

- 1.瞭解學習「統計學」的目的與價值。
- 2.學習判別各種原始資料的數值類型。
- 3.清楚分辨統計學的分類。

技能

- 1.能夠獨立判別資料的數值類型。
- 2.能夠自主判斷各種統計學的分類。

態度(情意)

- 1.意識到學習統計學對日常生活和未來生涯發展的重要性。
- 2.建立自己對統計學的新觀念。不再懼怕。

統計學是新數位化時代各領域科學發展的基礎。從消費者行為的預測、產品或服務消費量的預測、市場調查、產品開發、藥物開發、疫苗開發、疫情推估、醫療技術開發、管理效益評估、生產效益評估、行銷效益評估、服務效益評估、農作物品種改良、服務管理最佳化、手機 APP 使用模式、社群資訊分享、線上搜尋、旅遊行為、住宿選擇、餐飲消費行為、購物行為模式等，幾乎都是利用統計學為其發展的分析工具。

必須透過邏輯連貫、嚴謹、可靠、周延與客觀資料的蒐集、整理、歸納、分析、陳述、判讀、推論和預測，在不確定的狀況下建立科學性的定律和規則，歸納出通盤性的結論，進而在管理上提出最佳的決策和處理方式。

本章期望可以瞭解統計學的概況，認識數值資料的四種評量尺度，比較質化和量化資料的差異性與橫斷面和時間序列資料的差異性，瞭解不同的資料來源之特質，期待可以降低資料取得的錯誤，同時理解敘述性統計、推論性統計和應用統計的範疇。

「統計學」狹義定義是指數字化呈現物件、事實或資料的分布情況。個別資料與母體資料的差異。個別資料屬於單一基本單位的特定性質，所代表的一種數值資料，對其他基本單位不具有任何代表性意涵；母體資料屬於眾人多元化的性質，對特定基本單位所構成的母體具有一定程度的代表性意義。統計學可以視為一種從母體資料中擷取資訊的工具或方法。

「統計學」廣義定義是一系列資料蒐集、資料整理(組織)、分析、顯示(陳述、表現)、判讀(解釋資料)、推論分析結果和預測未來趨勢的嚴謹暨客觀程序，並藉科學的統計推論程序，在事先不瞭解(不確定)母體的分布情況下，透過抽樣蒐集到的樣本資料，藉由整理和分析獲得的資訊，推論母體資料的分布情況，進而發展適當決策(optimal decision)的一種科學領域(一門科學)。客觀解讀「數值」說服力和價值遠超過「道聽塗說」和「軼聞」。數值來源和蒐集方法必須具有代表性、客觀性和科學性，此數值才具有統計分析的意義和價值。

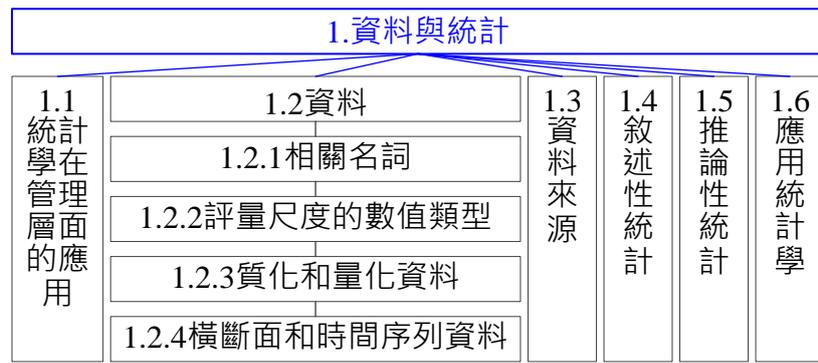
進行統計推論時，必須思考是否還有其他變數會影響目標變數，若有其他變數具有相當程度的影響，皆必須納入評估範圍。影響營利性遊樂區遊客重遊意願不應只是門票價格和服務品質兩個變數，其他旅遊體驗、動機、口碑、消費行為等變數也會影響重遊意願。

母體之統計數值的變異性是常態。母體或樣本的統計數值是眾多觀測值的代表數值，會隨著其他的因素而波動，一定幅度之內的波動是正常現象。例如：體重數值的變異(波動)、收入金額的變異、營業額的變異等。故，統計數值一般情況下不是一個恆定數值、定值或常數。

從統計學的基本概念，擴展到統計分析方法時，統計分析方法的步驟可以區分為蒐集：藉由實驗、觀察、調查或其他現有二手資料收集相關資料；彙整：通常利用表、圖或統計數值表現數值的分布狀況；解釋：陳述表、圖或統計數值所代表的意義；推論：透過統計理論對母體參數進行推估。

依據統計所推論的結果無法達到完全正確的程度。統計推論的結果成功機率不會達到 1.0000。

養成細心和周延學習的習慣與態度，認真處理數值統計的每一個階段，以解析數值所代表的真正意涵。



章節結構圖

1.1 統計學在管理層面的應用

會計(Accounting)

運用各種觀察和評量的方法，將商業交易及其相關事項紀錄、歸類或彙整，使資料具有意義和價值，並且加以分析和陳述特定企業體的財務狀況、營運情況暨現金流量，以提供管理階層數量化的財務資訊，協助其擬定決策或判斷。

財務(Finance)

銀行理財顧問利用各種統計資訊為投資客建議投資的標的。購買股票、房地產、基金、定存、黃金等。

行銷(Marketing)

產品開發、廣告效益、市場分析、銷售分析、通路評估、銷售預測等相關研究領域之運用。

百貨公司、量販店、便利商店常利用條碼掃描，蒐集消費者購買商品的資訊，進行量化統計分析，以提供行銷單位擬定行銷策略與依據。

生產(Production)、營運(Operation)和管理(Management)

工廠利用抽樣分析統計的方法，進行商品生產品質的控制，提高產品品質和可靠度。服務業利用抽樣問卷調查的方法，統計分析消費者的意見，以進行服務品質的監控。

經濟(Economics)

利用各種統計資訊建立經濟發展模式、作為預測未來經濟狀況或相關的發展方向之依據。國內官方統計資料以[中華民國統計資訊網](http://www.stat.gov.tw)主要資訊提供單位。

「統計學」在觀光與餐旅業的應用

營運資料的分析判讀：敘述性統計：表格和圖形法；敘述性統計：數量方法中央趨勢量、分散數量

市場調查與可行性影響評估：敘述性統計：表格和圖形法；敘述性統計：數量方法中央趨勢量、分散數量；推論性統計

消費金額、消費者數量、產品或服務販售數量和住房率的預測：簡單回歸分析、複回歸分析

彩券設計、博奕、摸彩活動設計：機率

產品品質或服務品質之檢驗：假設檢定、區間估計

市場區隔與目標市場選定：卡方分析、 t 分布、 t 檢定、 F 分布、Anova 分析

創新餐飲產品或創新服務模式開發最佳化：Anova 分析、簡單回歸分析、複回歸分析

行銷效益的評估：實驗法、 t 分布、 t 檢定、 F 分布、Anova 分析

學習「統計學」是在觀光與餐旅業從勞務職和面貌職，邁向現代化專業管理職和高薪職的基本要素。也是觀光與餐旅業產業升級的必備基本學能。強化學習的誘因與動機，克服數理學習的心理障礙，在數位化觀光餐旅世代創造出客觀、理性和價值。

1.2 資料

統計數值(數字)和**其相關的文字敘述**，可以歸類為欲統計分析的「**資料**」(data; data set)。在特定的時空和環境下，各種活動或自然現象母體中，有關人、事、時、地、物等，依據其相關**特徵**的敘述或記錄，皆可統稱為「資料」。故，資料可視為統計學欲分析的對象或素材。需要徹底了解「資料」的真正特質，才能理解統計分析的目的。

1.2.1 相關名詞

母體、族群、群體(Population or universe)或母群體(Parent population or parent universe)

在特定研究議題欲研究調查的可能對象，由一群具有某種共同特徵的基本單位所構成的群體(組合)。特定研究議題下，其所有基本單位(元素)的集合體。可以視為可能研究調查的對象之全部集合。母體中基本單位的數量一般利用 N 符號代表。

欲研究調查特定遊樂區遊客的旅遊動機，**在特定期間內蒞臨該旅遊區的遊客**，皆為該研究調查的**母體**。可否在高雄隨機選取消費者或學生，調查其將到墾丁國家公園的動機呢？

有限母體(Finite population)

母體所包含的基本單位數有一定數量，可以計數產生明確數值。一般母體特徵屬於**固定狀態**者。

無限母體(Infinite population)

母體所含的基本單位數數值很大，難以計量產生明確數值；非無窮大，但要實際計量，並不容易。一般母體特徵屬於**流動狀態**者。

基本單位(elementary units)、研究個體、個體(Individual)、元素(Element)

係指**母體**中可以被觀察或調查之**不可分割、個別份子或單獨(單一獨立)的個別成份**。在**社會科學(餐飲觀光休閒領域)**的母體中幾乎所有**基本單位**皆是指「**個人**」、「**消費者**」、「**遊客**」、「**居民**」、「**原住民**」、「**遊憩者**」、「**顧客**」、「**會員**」、「**員工**」、「**國人**」、「**主管**」為眾。基本單位除個人以外，亦有可能包含**服務業**、**零售業**、**製造業**、**市場**、**國家**、**甚至沒有生命的物體**皆可能是特定調查主題母體的**基本單位**。

樣本(Sample)

母體的**其中一部份基本單位**所**組成**的群體。在抽樣調查中，只蒐集及分析樣本的數值資料，然後根據樣本提供的資訊來瞭解和探索母體的情況，故**樣本**必須具有**代表性**。樣本中基本單位的數量一般利用 n 符號代表。

研究變數或變數(Variable)

在母體或樣本中每一個**基本單位**內，具有**多種對研究觀察或調查有價值(興趣)的特徵或特質**，稱為變數，經常使用大寫字母代表，例如：變數 X ，該變數的特定數值稱為變量(variate)，使用小寫字母代表 x 。例如：數量眾多之遊客(基本單位)的個人月收入(變數)，在特定 45000 元新台幣個人月收入數值(變量)、性

4/10/2024 7:36:36 PM 當您發現本教材錯誤時，盡速通知老師修改，教學才會進步。

別(變數)為女生(變量)、年齡(變數)為 25 歲(變量)、消費金額(變數)、旅遊同伴(變數)、家庭組成(變數)、動機(變數)、行為(變數)、認知(變數)、價值觀(變數)等。

觀察值或觀測值(Observations)

對特定**基本單位**之特定**變數**進行測量或觀察，而獲得的**測量值**。例如：特定遊客陳○○(基本單位)的年齡(變數)為 25 歲(觀察值)；消費金額(變數)新臺幣 2695 元(觀察值)；旅遊同伴(變數)為同學(觀察值)；性別(變數)是女生(觀察值)；學歷(變數)是研究所(觀察值)。

隨機抽樣(Probability sampling; Random sampling)

採用隨機(Randomness)做為樣本選取的控制機制。
隨機抽樣中每個基本單位被抽中的機率皆**相等**且獨立。

非隨機抽樣(Non-probability sampling; Nonrandom sampling)

以研究者、個人或條件判斷做為樣本選取的控制機制。
無法說明每一個基本單位被抽取的機率。

母數、母體參數(population parameter)、介量或參數(Parameter)

代表**母體**中特定特徵、表徵、屬性或變數的數值(測量值)。即研究調查欲獲得的核心數值。
母體特定屬性的平均值(mean, M)、標準偏差(standard deviation, SD)、變異數、比例等。

統計值、統計量(statistics)、樣本統計值、樣本統計量(sample statistics)或估計值(Estimate)

由**樣本**獲得的特定特徵、表徵、屬性或變數的數值(測量值)，用以推估母體參數。
樣本特定屬性的平均值(mean, M)、標準(偏)差(standard deviation, SD)、變異數、比例等。

估計誤差

利用樣本的統計值估算母體的參數時，難免在樣本的統計量和母體的參數會產生差距。差距分為抽樣誤差和非抽樣誤差兩種。

抽樣誤差(Sampling error)

The degree of error to be expected for a given sample design (Babbie, 2001).

因抽樣時樣本可能會偏離母體，其間的差距，可用統計方法估算其差距大小。抽樣經分析獲得的樣本統計值(statistic)與母體母數(parameter)的差距。

樣本中若包含某些特殊屬性的基本單位，會破壞樣本的代表性。可能是由碰巧(chance)、抽樣方法(sampling method)、抽樣架構及推論方法所致。通常可以透過增加樣本數量和執行隨機抽樣程序，降低抽樣誤差。

非抽樣誤差(Non-sampling error)

因抽樣誤差以外的因素，所導致的差距。主要是由人為因素所導致，可能是資料填寫不實、遺漏、錯估，或資料整理分析時輸入錯誤、計算錯誤、反向問題數值未導正等。

非抽樣誤差的大小無法精確估算，操作上會假設非抽樣誤差很小，可忽略不計。惟實際操作層面若研究者無經驗、花錢誤事、無心、無知、粗心、無人教導提醒、沒有重複確認、請他人代工等，可能會使非抽樣誤差變大，甚至於高過抽樣誤差，成為研究的主要誤差來源，致使研究結論的推估完全錯誤。

甚至於研究者或其團隊事後亦完全不知、無知或沒有察覺。透過細心、專心和用心，期望可以降低非抽樣誤差。

範例 1.1 欲調查本大學全部學生的飲食行為，設計採用隨機方式抽選其中 300 位學生進行問卷填答蒐集資料。請問何者是母體？何者是樣本？何者是研究變數？

答案：母體：本大學全部學生；樣本：隨機方式抽選的 300 位學生；研究變數：飲食行為

作業 1.1 導師想瞭解導師班級 58 位學生中，每位學生平均每日花在看書的時間分布情況。從全班同學隨機抽出 17 位學生為其樣本分析。試問：(a)母體？(b)所要研究的變數？(c)樣本？

1.2.2 評量尺度數值類型

在測量特定屬性或變數時，回答選項有四種數值尺度或量表水準(levels of scales)可以使用，包括名目尺度(Nominal scale)、順序尺度(Ordinal scale)、等距尺度(Interval scale)和比例尺度(Ratio scale)。在評量歸類屬性或變數時，必須考慮其回答選項的設計，不同回答選項設計方式，其歸屬數值尺度可能不相同。

1.2.2.1 名目尺度

名目尺度、名義尺度(Nominal scale)或類別尺度(Categorical scale)是四個量表水準中，最基本的一種類型。此類型的問題，讓受訪者回答時，只需要提供一些名稱(names)、符號(label)、類型(category)等型態的回答選項。此型態回答選項中沒有包含任何具有強度水準之內涵，因此無法給予回答選項之間有任何的順序型態。

名目尺度主要之目的即是鑑別(identify)數值的特性、組成或分布。評量尺度中數值化程度最低者，僅能夠呈現資料中各類別的比例，資料可量化統計分析程度最低。數值型態屬於非連續性變數(Non-continuous variable)和離散變數(discrete variable)，名目尺度(Nominal scale)評量題目收集到的數值(代碼)中，若具有小數點後的數值，不具任何意義。

名目尺度回答選項之間需要具備互斥性和周延性。獲得的資料才具有可靠度。

- a. **互斥性(mutually exclusive)**或**選項屬性互斥(mutually exclusive)**：在同一個項目(題目、問題)中，回答的(類別)選項之間屬性相互排斥、無相互重疊之處。不要讓受訪者覺得此題目的回答選項有複選題之感覺。
- b. **周延性(exhaustive)**或**選項涵蓋完整**：不讓受訪者覺得問題之答案不在選項中，設計回答選項必須周延(exhaustiveness; inclusiveness)，考慮各種會出現的答案選項，皆必須列入回答選項中。

在蒐集名目尺度資料的過程中，設計回答選項若沒有達到互斥性和周延性，其所蒐集獲得的資料並無法具體地呈現資料的分布，資料就沒有分析的價值。

名目尺度(Nominal scale)可區分為

標記(Label)

記載事物的工具，而非進行數量分析；數字不代表屬性的多少、高低、大小，僅是方便紀錄的代號。

例如：飛機航班編號、車牌號碼、桌號、菜單(menu)內菜色(item)編號。

類別(Category)

使用數字代表物體歸屬的類別。將被測量的對象分成幾種類別，同一個類別內所有樣本皆賦予一個相同的數字代表。

例如：性別、職業、宗教類型、血型、國籍、居住縣市、交通工具或顏色。

特性：可分類別或分群體。沒有前後順序、沒有大小和沒有距離。類別(項目)定義要明確一致，以免造成受訪者認知誤差。例如：收入：高、中、低；高雄居住地：海邊、山岳、市區、農村。選項無明確定義，評量效益差。

數值處理：只能將各類別(回答選項)出現的頻率利用百分比(%)呈現變數分布型態，進行統計分析。可以計算此類資料的眾數。

例如：請問您這學期最喜歡的是哪一門課程？ 統計學 會計學 經濟學 營養學 體育 其他_____

1.2.2.2 順序尺度

順序尺度、次序尺度、序位尺度或等級尺度(Ordinal scale)允許受訪者依據回答選項之間，可以進行相對大小的比較，**選擇一個與其感受相同程度選項**。因此，於設計回答選項時，即呈現出具有層次選項的模式。如同**非常大/非常小、非常高/非常低、非常頻繁/非常不頻繁、非常重要/非常不重要、非常同意/非常不同意**等類型的回答序列選項。

具有**名目(類別)尺度**的所有特性(**互斥性**和**周延性**)。再加上回答選項之間透過相互的**比較**有大小、優劣、同意程度、前後或順序的關係存在。實務上在評量「年齡層分布」、「成績名次」、「收入階層高低」、「教育程度高低」等變數皆是屬於順序尺度。

順序尺度數值型態屬於**非連續性變數**(Non-continuous variable)和**離散變數**(discrete variable)。**順序尺度數值特性**可比大小及前後，但前後距離不等、不等距。例如 $5 > 4$ ，但 $5 - 4 \neq 1$ 。**順序尺度數值處理上**只能將各層次選項出現的頻率利用百分比(%)呈現變數分布型態，進行統計分析。

順序尺度設計基本範例：扣除最前和最後選項，中間選項區間寬度盡量一致，建構相同比較基礎的概念。區間寬度小時，方便後續合併組別彈性選擇空間。

請問您的年齡？16-20 21-25 26-30 31-35 36-40 41-50 51 歲以上

請問您此次同遊人數？獨自前往 2 人 3 4 5 6 人(含以上)

請問您過去一年來到花蓮農場旅遊的次數？1 次 2 3 4 次(含)以上

請問您此次到花蓮農場旅遊，前後預計停留多少時間？不到一天 二天一夜 三天二夜 四天以上

1.2.2.3 等距尺度

在等距尺度、區間尺度或間距尺度(Interval scale)中，具有**名目(類別)**和**順序**尺度特性，並可**比較差距**。回答選項之間**前後距離相等**，間隔(Interval)可以用來加減，但沒有倍數關係，所以**沒有絕對 0 點**(logical zero points、原始點、基準點 natural origin、natural zero、rationale zero point)、**有相對 0 點**，0 點(原始點或基準點)可設定(被定義)於**任一點**(數值)。等距尺度無法做倍數計算，例如 $4 \neq 2 \times 2$ (第四個選項 \neq 第二個選項 $\times 2$)， 40°C 不能詮釋為 20°C 的 2 倍。

等距尺度的數值若評量為負值(-)時，因此負值是被定義產生，故該評量結果具有意涵存在。評價尺度(Rating scale)和李克特量表(Likert scale)中，若回答選項設計為「非常同意」、「同意」、「普通(不確定)」、「不同意」、「非常不同意」~~分別依序給予 5、4、3、2、1 分代表~~，原本只能視為順序尺度

(Ordinal scale)，但在社會科學的研究中，研究者將其間差距視為等距(equal interval)，故評價尺度(Rating scale)和李克特量表(Likert scale)實務運用上被視為等距尺度。攝氏溫度(°C)和華氏溫度(°F)，僅具有相對的0度，沒有絕對的0度，故攝氏溫度(°C)和華氏溫度(°F)獲得的數值類型屬於等距尺度。

等距尺度數值處理上可以進行百分比、平均值和標準(偏差)的運算、李克特量表(Likert scale)的數值可以具有平均值(Mean, *M*)與標準(偏差)(Standard deviation, *SD*)。

等距尺度範例(回答選項之間強調距離等距特徵；或標示5、4、3、2和1分代表)：請問您此次到花蓮農場旅遊，整體的滿意度？

非常滿意 滿意 普通 不滿意 非常不滿意

1.2.2.4 比例尺度

在比例尺度、等比尺度、比率尺度或比值尺度(Ratio scale)中，具有名目(類別)、等級、等距尺度特性，且可以做倍數計算，具有可乘性(multiplicativity)。例如， $4 = 2 \times 2$ 。有距離，具有絕對的零點(零的數值有其意義，沒有負值或負值不存在 natural origin)。0點、基準點、原始點位置(數值)固定不會變動。

比例尺度的數值若評量為負值時，因數值屬於絕對性質，故該評量結果不具有任何意涵存在。因此，比例尺度絕對不會出現負值。

例如：收入數值 21,000 元、年齡數值 20 歲、身高 160.3 cm、體重 56.2 kg、價格 25 元、費用 2,000 元、年銷售量(額)2,523,781 元。絕對溫度 500 K，絕對溫度(K) = 攝氏溫度(°C) + 273.16 度，在自然界中絕對溫度不會有負值，所以絕對溫度就是擁有絕對 0 點，歸類為比例尺度。

數值處理：所有基本運算皆可行，具有加減乘除四則運算特性。數值統計分析處理的效益與價值最高。

比例尺度設計範例：

請問您此次同遊人數？(不包括自己)_____人

請問您此次到花蓮農場四天三夜旅遊，一共花費多少錢在購物上？(新台幣)_____元

名目尺度(Nominal scale)和順序尺度(Ordinal scale)歸屬於**非計量尺度**(nonmetric scale; nonmetric measurement scales)。非計量尺度可以在細分為 Binary、Nominal、Ordered、Ranked 等四種。等距尺度(Interval scale)和比例尺度(Ratio scale)歸屬於**計量尺度**(metric scale; metric measurement scales)。

1.2.2.5 問卷設計中數值類型敘述範例

1.2.2.5.1 旅遊和遊憩行為

變 項 名 稱	數值類型	答 案 項 內 容	引用文獻
請問您此次同遊人數？	順序尺度	<input type="checkbox"/> 獨自前往 <input type="checkbox"/> 二人 <input type="checkbox"/> 三 <input type="checkbox"/> 四 <input type="checkbox"/> 五 <input type="checkbox"/> 六人(含以上)	Chou, 1999
請問您此次同遊人數？(不包括自己)	比例尺度	_____人	Lee et al., 2004
請問您過去一年來到花蓮農場旅遊的次數？	順序尺度	<input type="checkbox"/> 一次 <input type="checkbox"/> 二 <input type="checkbox"/> 三 <input type="checkbox"/> 四次(含)以上	Tsai & Chou, 2002

變項名稱	數值類型	答案項內容	引用文獻
請問您此次到花蓮農場旅遊，前後大約停留多少時間？	順序尺度	<input type="checkbox"/> 不到一天 <input type="checkbox"/> 二天一夜 <input type="checkbox"/> 三天二夜 <input type="checkbox"/> 四天以上	自行歸納設計
請問您此次到花蓮旅遊，一共花費多少錢？(新台幣)	比例尺度	_____元	Tsai & Chou, 2002
請問您此次到花蓮農場，整體的滿意度？	順序尺度 等距尺度	<input type="checkbox"/> 非常不滿意 <input type="checkbox"/> 不滿意 <input type="checkbox"/> 普通 <input type="checkbox"/> 滿意 <input type="checkbox"/> 非常滿意	自行歸納設計
請問此次誰與您一起到花蓮農場？	名目尺度	<input type="checkbox"/> 同學 <input type="checkbox"/> 朋友 <input type="checkbox"/> 家人 <input type="checkbox"/> 同事 <input type="checkbox"/> 其他_____	Tsai & Chou, 2002

1.2.2.5.2 社會經濟背景

變項名稱	數值類型	答案項內容	引用文獻
性別	名目尺度	<input type="checkbox"/> 男 <input type="checkbox"/> 女	Chou, 1999
年齡	比例尺度	_____歲	Lee et al., 2004
年齡	順序尺度	<input type="checkbox"/> 20歲以下 <input type="checkbox"/> 21~30 <input type="checkbox"/> 31~40 <input type="checkbox"/> 41~50 <input type="checkbox"/> 51~60 <input type="checkbox"/> 61歲以上	Tsai & Chou, 2002
教育程度	順序尺度	<input type="checkbox"/> 國小及以下 <input type="checkbox"/> 國中 <input type="checkbox"/> 高中職 <input type="checkbox"/> 專科 <input type="checkbox"/> 大學 <input type="checkbox"/> 研究所	自行歸納設計
職業	名目尺度	<input type="checkbox"/> 服務業 <input type="checkbox"/> 軍公教 <input type="checkbox"/> 金融業 <input type="checkbox"/> 營造業 <input type="checkbox"/> 電子業 <input type="checkbox"/> 大眾傳播業 <input type="checkbox"/> 製造業 <input type="checkbox"/> 運輸業 <input type="checkbox"/> 通信業 <input type="checkbox"/> 商業 <input type="checkbox"/> 自由業 <input type="checkbox"/> 家庭主婦 <input type="checkbox"/> 學生 <input type="checkbox"/> 待業中 <input type="checkbox"/> 退休 <input type="checkbox"/> 其他_____	自行歸納設計
個人平均月收入(新台幣：元)	順序尺度	<input type="checkbox"/> 20,000以下 <input type="checkbox"/> 20,001~30,000 <input type="checkbox"/> 30,001~40,000 <input type="checkbox"/> 40,001~50,000 <input type="checkbox"/> 50,001~60,000 <input type="checkbox"/> 60,000以上	Tsai & Chou, 2002
家庭月收入(新台幣：元)	順序尺度	<input type="checkbox"/> 20,000以下 <input type="checkbox"/> 20,001~30,000 <input type="checkbox"/> 30,001~40,000 <input type="checkbox"/> 40,001~50,000 <input type="checkbox"/> 50,001~60,000 <input type="checkbox"/> 60,000以上	自行歸納設計
目前居住地	名目尺度	<input type="checkbox"/> 台北市 <input type="checkbox"/> 新北市 <input type="checkbox"/> 基隆市 <input type="checkbox"/> 桃園縣 <input type="checkbox"/> 新竹市 <input type="checkbox"/> 新竹縣 <input type="checkbox"/> 苗栗縣 <input type="checkbox"/> 台中市 <input type="checkbox"/> 彰化縣 <input type="checkbox"/> 南投縣 <input type="checkbox"/> 雲林縣 <input type="checkbox"/> 嘉義縣 <input type="checkbox"/> 嘉義市 <input type="checkbox"/> 台南市 <input type="checkbox"/> 高雄市 <input type="checkbox"/> 屏東縣 <input type="checkbox"/> 宜蘭縣 <input type="checkbox"/> 花蓮縣 <input type="checkbox"/> 台東縣 <input type="checkbox"/> 澎湖縣 <input type="checkbox"/> 金門縣 <input type="checkbox"/> 連江縣 <input type="checkbox"/> 其他_____	Tsai & Chou, 2002

練習 1.1 「攝氏溫度」在量度尺度分類中屬於下列哪一種？(A)名目尺度(nominal scale)；(B)順序尺度(ordinal scale)；(C)等距尺度(interval scale)；(D)比例尺度(ratio scale)(2007 初等考試統計學大意)

練習 1.2 溫度使屬於下列哪一種變數型態？(A)比例尺次(ratio scale)；(B)等距尺度(interval scale)；(C)順序尺度(ordinal scale)；(D)比例尺度(ratio scale)或等距尺度(interval scale)

練習 1.3 「足球員的球衣號碼」在量度尺度分類中屬於下列哪一種？(A)名目尺度(nominal scale)；(B)順序尺度(ordinal scale)；(C)等距尺度(interval scale)；(D)比例尺度(ratio scale)(2007 初等考試統計學大意)

練習 1.4 下列哪一組是分類為質化資料(qualitative data)？(A)等距與順序資料；(B)比率與順序資料；(C)等距與名目資料；(D)名目與順序資料(99 年公務人員初等考試統計學大意)

1.2.3 質化和量化資料

蒐集的資料可以透過其數值化的程度區分為質化資料和量化資料。

質化資料、質的資料、質性資料、定性資料或屬性資料(Qualitative data)

調查變數的特徵屬於非數值型態。需利用一組的標記、類別、性質或名稱以區別每個基本單位的特徵、屬性。無法以數值表示的統計資料，如航班編號、性別、學歷、旅遊同伴、頭髮顏色、宗教等。通常名目尺度和順序尺度會被歸類為質化資料。

若變數屬於質化資料，統計上對於每一個種類下的數量或比例，為一般的描述方式。

可能是數值型態或非數值型態。

量化資料、數量資料、量的資料、量性資料、定量資料或屬量資料(Quantitative data)

可以使用數值表示的統計資料，如銷售量、銷售金額、收入、支出、身高、體重、年齡、小孩數量等。通常等距尺度和比例尺度會被歸類為量化資料。

調查變數的特徵屬於數值型態。

在典型的量化資料中依據數值分布是否為連續性數值型態，可以區分為間斷性資料(discrete data)和連續性資料(continuous data)兩種類型。

間斷性資料、離散性資料(Discrete data)(計數資料)

數值的數量是有限或可以計數者，一般僅用特定的數值代表，而且數值之間具有固定的「間隔」存在。例如：房間數量、小孩人數、蛋數、特定購餐車道一小時內通過汽車數量。

以點計方式獲得的資料。原始間斷性資料小數點後面的數值沒有意義。

連續性資料(Continuous or numerical data)(計量資料)

數值的數量是無限，屬於連續式的量表型態，在數值區間中沒有間隙(gap)、中斷(interruptions)或跳過(jumps)。小數點後面的數值，有其明確意義存在。例如：高壓鍋中壓力數、蘋果重量、用餐時間、公尺數、公升數、公斤數、熱量。

以測量方式獲得的資料。

範例 1.2 下列特定餐廳中相關資料，何者屬於質性資料？何者屬於計數資料(數值僅能表達至個位數)？何者屬於計量資料(數值可表達至小數者)？

- | | |
|------------------|--------------------|
| (a) 特定員工使用的交通工具 | (b) 昨天上班的員工人數 |
| (c) 每天每位員工平均上班時數 | (d) 販售數量最多的餐點 |
| (e) 平均每日營業額 | (f) 昨天販售特定餐點的數量(盤) |

答案：質性資料：(a)特定員工使用的交通工具、(d)販售數量最多的餐點；計數資料：(b)昨天上班的員工人數、(f)昨天販售特定餐點的數量(盤)；計量資料：(c)每天每位員工平均上班時數、(e)平均每日營業額

1.2.4 橫斷面和時間序列資料

依據收集資料的時間點或區間之數量，可以將資料區分為橫向面資料、時間序列資料和縱橫資料三類。

橫向面資料、橫斷面資料(cross-sectional data)

同一時間或在同一**特定時間區間**內蒐集到的資料。一般指的是在此特定時間點，欲研究觀察之變數的觀察值。例如：欲評量本班全班同學體重的分布狀態，於特定一個時間全班統一帶到健康中心依序測量每位同學的體重資料，如此獲得的資料即屬於橫向面資料。

時間序列資料、時間數列資料(time series data)

在一段時間區間內，連續對特定變數的觀察值。主要是希望呈現不同時間點，該特定變數的觀察值之變化情況。例如：欲評量特定同學體重大一的變化情況，從入學後第一個月即到健康中心測量體重，依序第二個月也到健康中心測量體重，此此進行量測 12 個月，此 12 個月量測的體重資料即屬於時間序列資料。飯店過去三年中每個月住房率的資料，亦屬於時間序列資料。

縱橫資料(panel data)

同時包含橫向面資料和時間序列資料。例如：欲評量本班全班同學體重大一一整年的分布變化狀態，於開學九月特定一個時間全班統一帶到健康中心依序測量每位同學的體重資料，十月亦約好一個時間全班統一帶到健康中心依序測量每位同學的體重資料，依序進行 12 個月，如此獲得的資料即屬於縱橫資料。

依據調查分析目的需求，對每個基本單位中蒐集的變數數量，可以將獲得的資料區分為單變量資料、雙變量資料和多變量資料三類。

單變量資料(univariate)

針對**每個基本單位**只記錄一個(種、類、變數)資料，所以同一個基本單位只針對**一個變數進行評量**。主要目的進行該變數分布的敘述性統計。例如：下表中旅館名稱代碼，代表獨立各別的基本單位(特定旅館)；每一個旅館皆有一個對應的一個平均房價資料，即每一個基本單位(旅館)都有其對應的平均房價資料。

旅館名稱	平均房價
A	NT\$3562
B	NT\$3962
C	NT\$3462
D	NT\$4662
E	NT\$9356
F	NT\$8562
G	NT\$7562

雙變量資料(bivariate)

針對**每個基本單位**只記錄**兩個(種、類或變數)資料**，所以同一個基本單位同時只針對**兩個變數進行評量**。主要目的除了進行兩個變數分布的敘述性統計以外，也企圖理解兩個變數之間的關係。兩個變數之間具有配對關係。例如：下表中旅館名稱代碼，代表獨立各別的基本單位(特定旅館)；每一個旅館皆有兩

個對應的資料(分別有平均房價和房間數量)，即每一個基本單位(旅館)都有其對應的「平均房價」和「房間數量」兩個資料。

旅館名稱	平均房價	房間數量
A	NT\$3562	250
B	NT\$3962	265
C	NT\$3462	350
D	NT\$4662	450
E	NT\$9356	225
F	NT\$8562	352
G	NT\$7562	450

多變量資料(multivariate)

針對**每個基本單位**只記錄**三個或三個(種、類或變數)以上資料**，所以同一個基本單位同時針對**三個或三個以上變數進行評量**。主要目的除了進行多個變數分布的敘述性統計以外，也企圖理解多個變數之間的關係。變數之間具有配對關係。例如：下表中旅館名稱代碼，代表獨立各別的基本單位(特定旅館)；每一個旅館皆有多個對應的資料(分別有平均房價、房間數量、品質、類形、前一個月住房率)，即每一個基本單位(旅館)都有其對應的「平均房價」、「房間數量」、「品質」、「類型」、「前一個月住房率」多個資料。

旅館名稱	平均房價	房間數量	品質	類型	前一個月住房率
A	NT\$3562	250	好(Good)	都會型旅館(City hotel)	80%
B	NT\$3962	265	好(Good)	渡假型旅館(Resort)	70%
C	NT\$3462	350	好(Good)	渡假型旅館(Resort)	76%
D	NT\$4662	450	尚可(Fair)	都會型旅館(City hotel)	75%
E	NT\$9356	225	好(Good)	一般型旅館(Hotel)	95%
F	NT\$8562	352	尚可(Fair)	渡假型旅館(Resort)	60%
G	NT\$7562	450	優等(Excellent)	渡假型旅館(Resort)	55%

1.3 資料來源

依據研究或調查的目的，決定需要蒐集哪些類型或來源的資料。依據資料的來源(Data sources)，可以區分為原始資料和次級資料。

次級資料(secondary data)、間接資料、現有資料 Existing sources

所需的資料已存在公司、團體或組織的資料庫中，不是為了此次調查分析目的而特定蒐集的資料，如過去的銷售記錄、過去會員的消費記錄等。

原始資料(primary data)、直接資料、統計分析研究獲得的資料 Statistical studies

所需的資料無法從現存的資料庫(來源)中獲得，必須依據此次的調查分析目的，利用研究方法取得代表性的資料，可以分為：實驗性(experiment)研究、調查(survey)與觀察性(observation)研究。

實驗性研究是指控制環境中操作特定變數，獲得研究想要蒐集的(觀察)變數之觀察值。主要是透過實驗性的研究設計，以瞭解操作變數與觀察變數之間的互動關係。

調查性研究是對影響母體的各种因素不做控制操作，進行資料蒐集的方法。經常使用訪談法或問卷調查法，針對研究對象蒐集特定議題的意見或反應。

觀察性研究是不控制或影響環境中任何變數，利用觀察記錄的方式，針對研究對象蒐集特定議題的相關資料。之後，針對記錄進行分析，以歸納出研究結論。

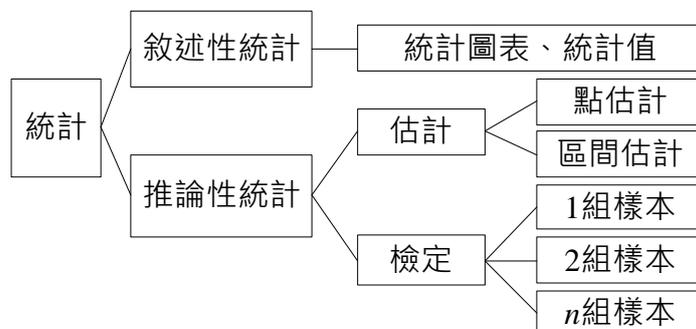
資料取得之錯誤(Data Acquisition Errors)

在統計研究中，研究者應該必須知道**可能獲得錯誤的資料**，若**無法排除錯誤的資料**，恐會造成**統計推論的錯誤**。故在資料蒐集的過程中，必須特別注意如何排除錯誤的資料。

統計學依據其主要運用目的可分為：**敘述性統計**(descriptive statistics)、**推論性統計**(inferential statistics)和**試驗設計**(experimental design)。

試驗設計、實驗設計(experimental design; design of experiments)

設計一種特殊的環境、情境，非自然環境(情境)，以驗證研究假設是否成立。運用試驗過程(資料蒐集)中之重複性與隨機性，使研究變數(實驗變數、自變數)以外之其他變數(含已知與未知)對觀察變數或觀測變數(依變數)的影響相互抵銷，以單純化評量研究變數(實驗變數、自變數)對觀測變數(依變數)的影響效益，以提升分析結果的準確性。以檢驗實驗研究假設中，研究變數(實驗變數、自變數)對觀測變數(依變數)之關係。



1.4 敘述性統計

敘述性統計(學)、敘述統計(學)(Descriptive statistics)、描述統計學

對於母體或樣本中特定變數的大量資料(觀測值)，可透過數值、統計圖或統計表將大量的資料，清楚和直接的呈現資料所代表的訊息或意涵。充分利用各種數值、圖、表、線條的呈現方式，表達出母體或樣本變數資料的分布情況，盡可能達到「一目了然」的水準。通常敘述性統計僅呈現樣本或母體的資料，不對全部母體的資料進行推論，即不將樣本資料分析結果之意涵，推展至更大的範疇。

敘述母體或樣本**特定變數分布狀態**的一種統計方法。

1.5 推論性統計

推論性統計(學)、**推論統計**(學)(Inferential statistics)、**統計推論**(學)(Statistical inference)或**歸納統計學**(Inductive statistics)

在絕大部分的情況下，科學性研究調查皆受制於人力、物力、財力、時間等資源的限制，無法對母體中所有的基本單位進行評量或調查(普查)，只能夠透過抽選母體中的部分基本單位進行評量或調查。

依據**樣本中特定變數**的資料(觀測值)之分析結果，對**母體該特定變數**的分布情況進行的一種推測(推論)、預測或估計程序。其所歸納出的結論，會超出樣本的資料範圍。

利用機率論的原理，進行**點估計**(point estimation)、**區間估計**(interval estimation)和**假設檢定**(testing hypothesis)三個階段，建立統計的模式(model)。

統計學依據其應用層次可分為：數理統計、數理統計學(mathematical statistics)和應用統計、應用統計學(applicative statistics; applied statistics)。

數理統計(學)、理論統計(學)、理論統計

運用數學相關理論闡述統計方法的原理，推導各種統計公式，研究各種統計推論的方法與程序為整個統計學的理论基礎。利用數學理論推論統計方法的原理，驗證各種統計公式的來源，研究發展統計推論的方法而為統計學的理论基礎者。故數理統計學主要闡明各種統計方法的原理。

範例 1.3 下列各種陳述中，哪些屬於敘述性統計？哪些屬於推論性統計？

- (a)全球暖化致國際小麥價格一直飆漲，預計下一季國內麵包產品售價會上漲 3 %。
- (b)大芳咖啡館上一季營業額大幅提昇，其原因主要是聘請到一位漂亮又負責的服務員 - 小玉。
- (c)近期台幣升值，導致進口成本降低，故進口車售價預計於下一季會調降 5 %。
- (d)小棋飲料店依據以往的銷售記錄顯示，購買珍珠奶茶的女性比例高達 70 %。

答案：敘述性統計：(b)、(d)；推論性統計：(a)、(c)

1.6 應用統計學

運用統計相關理论和統計方法在各種科學領域、學術及管理上，以解決實際面臨的統計問題者，可歸類為**應用統計學(Applicative statistics)**。應用於各領域，如：教育統計、政府統計、人口統計、經濟統計、生物統計、社會統計、國際貿易進出口統計...等。故應用統計學主要在闡明統計方法在各種學術研究及行政管理上之應用。

在實際的驗證過程中，數值的特性皆非單一性，因此，在分析實際的數值時，必須依據統計假設的條件下，預先建立模式，此模式中的參數尚未知道。透過樣本數值的資料，進行參數推估與檢定確認模式的參數值。

母數統計學、有母數統計學(parametric statistics)或母數統計法(parametric method)

母體欲觀察研究之變數屬性或特徵分布屬於**常態分布**或**趨近於常態**情況，分析研究的樣本資料時，在統計的假設條件下，事先建立分析模式，此分析模式中皆會有未知的參數(parameter)。透過樣本的資料，進行參數(parameter)的估計與檢定確定分析模式中的參數數值，利用此模式進行分析或預測。

例如： t 檢定、變異數分析(analysis of variance; ANOVA)、相關分析(correlation analysis)、迴歸分析(regression analysis)、時間序列分析(time series analysis)等。

無母數統計學(nonparametric statistics; non-parametric statistics)或無母數統計法(nonparametric method)

在樣本資料中無法滿足部分統計的假設條件，無法建立具有參數的統計分析模式之推論性統計方法。適合使用於母體欲觀察研究之特定變數的屬性或特徵分布不為常態也不易轉換為常態、母體分布情況不明(不確定)、樣本數量 n 小、類別或順序變項。運用樣本資料中的大小、順序或等級之特性，進行統計推論之方法。無母數統計中無須事先假設母體具有特定分布狀況，故推論過程也不限於任何型態的母體，皆可進行統計推論。

練習 1.5 台灣觀光局欲蒐集入境觀光客的資料，調查對象為桃園國際機場和小港國際機場入境的觀光客，利用問卷詢問下列問題：

- a.這是您第幾次到訪台灣：第 1 次 第 2 次 第 3 次 第 4 次 第 5 次以上
- b.此次到訪台灣的目的：觀光 商務 訪友 學術交流 醫療 其他

c.計畫住宿地點：旅館 親友家 招待所 宿舍 其他

d.預計停留台灣幾天？_____天

試評估

- 1.此調查的母體？
- 2.問卷調查入境觀光客是否是最佳方法？
- 3.上述四個問題的數值型態為？

作業 1.2 評量名目尺度或名義尺度(Nominal scale)的變數時，為什麼需要具備周延性(exhaustive)和互斥性(mutually exclusive)特性？以單獨 Microsoft word 撰寫答案，word 電子檔案繳交至數位學習平台，作業名稱：名目尺度特性，Word 檔案主檔名：姓名學號。繳交截止日期時間：依據平台設定為準。作業檔案來不及繳交時，不用 e-mail 給老師，留著以後有機會再交。

練習 1.6 請說明下列項目屬於哪一種測量尺度？可以依據不同回答選項的設計方式，歸類屬於測量尺度的類型。(A)學生這學期修課學分數量；(B)第一次統計學小考成績；(C)學生原始居住地；(D)學生到校上課的交通工具；(E)學生身上攜帶現金數量；(F)氣溫；(G)學生學號；(H)遊客年收入；(I)本班學生今天到校上課的前後順序。(提醒：回答方式會影響到數值類型)

練習 1.7 何謂敘述性統計(Descriptive statistics)？何謂推論性統計(Inferential Statistics)？請說明它們的內容，並討論它們兩者在統計學上的意義。(暨南公共行政)

練習 1.8 針對一組樣本欲代表其特徵的名詞稱為：(A)參數(Parameters)；(B)統計值(Statistics)；(C)機率(Probabilities)；(D)樣本(Samples)；(E)以上皆非。

練習 1.9 下列哪一個屬於質化變數(qualitative variable)？(A)製作珍珠奶茶的時間；(B)最常購買珍珠奶茶的連鎖店名；(C)哈雷機車的重量；(D)製作珍珠奶茶冰塊平均重量；(E)以上皆非。

練習 1.10 問卷調查民眾贊不贊成晶圓廠開放赴大陸投資，答項為非常贊成，贊成，中立，不贊成，非常不贊成等五項，請問這種資料的量測尺度是屬(a)類別；(b)名目；(c)順序；(d)等距；(e)比率。(成大國企所)

練習 1.11 導師想瞭解導師班級 58 位學生擁有機車的比率。現今從全班同學隨機抽出 10 位學生為其樣本，結果發現有 2 位學生有機車代步，由於機車的便利性高，預計下學期再調查時，可能會有 40 %的同學擁有機車。試問：(a)母體？；(b)所要研究的變數？；(c)樣本？(d)哪一部份依據敘述性統計？哪一部份依據推論性統計？

練習 1.12 調查觀光系一年級 A 班上 15 位學生的體重，並依據此所獲得的資料，做以下的陳述：全校同學平均體重為 62 公斤，此是下列哪一個名詞的範例？(A)敘述統計(descriptive statistics)；(B)推論統計(inferential statistics)；(C)參數(parameter)；(D)母體(population)

練習 1.13 請說明基本單位、變數和觀測值的定義？

練習 1.14 請將下列變數填入適當的分類表格中。依據自己的觀察與評估填寫。例如：薪資屬於連續性量化變數，同時也是屬於比例尺度。

- | | |
|-------|--------------|
| a. 薪資 | j. 學生出生地 |
| b. 性別 | k. 學生的 IQ 分數 |

- | | |
|-------------------|--------------|
| c. 去年郵輪停靠高雄港數量 | l. A 生上週唸書時間 |
| d. 咖啡飲料偏好 | m. 兩個城市的距離 |
| e. 溫度 | n. 旅遊行程的天數 |
| f. 統計學學期成績 | o. 每天運動時間 |
| g. 學生成績在班上排名 | p. C 員工的工作年資 |
| h. 此次前往墾丁希望投宿旅館順序 | q. 餐飲部門員工數量 |
| i. 家中電腦數量 | |

	間斷變數	連續變數
質化資料		
量化資料		薪資

	間斷變數	連續變數
名目尺度		
順序尺度		
等距尺度		
比例尺度		薪資

練習 1.15 下表是特定縣市國際觀光飯店的資料。

平均房價	房間數量	品質	類型
NT\$3562	250	好(Good)	都會型旅館(City hotel)
NT\$3962	265	好(Good)	渡假型旅館(Resort)
NT\$3462	350	好(Good)	渡假型旅館(Resort)
NT\$4662	450	尚可(Fair)	都會型旅館(City hotel)
NT\$9356	225	好(Good)	一般型旅館(Hotel)
NT\$8562	352	尚可(Fair)	渡假型旅館(Resort)
NT\$7562	450	優等(Excellent)	渡假型旅館(Resort)

- 上述資料的基本單位？
- 哪些是屬於量化變數？哪些屬於質化變數？
- 若有屬於質化變數時，其分別屬於名目尺度或順序尺度？請分別標示。
- 若有屬於量化變數時，其分別屬於等距尺度或比例尺度？請分別標示。
- 上述資料屬於橫斷面資料或時間序列資料？

建議閱讀資料

葉偉文譯(2006)。統計改變了世界，天下文化。

Salsburg, D. (2001). *The lady tasting tea: How statistics revolutionized science in the twentieth century*. New York: Henry Holt and Company.

討論議題

- 師生非同步教學討論議題：學習統計學價值

第一回合請於 D+3 日晚上 12 點以前從「議題討論」區【張貼】標題：「統計學價值」，本文：請依據階段自己的興趣與規劃，在未來的職業生涯發展中，最有興趣運用統計學在工作領域的情境與需求，請具體陳述亦可舉例說明(50 個字以上詮釋)。

期望可以透過議題討論資訊的交流，相互激勵，提升學習效益。待有 40 篇第一回合【張貼】回應後或第一回合【張貼】時間結束後，檢視其他同學的回應內容。自己靜下心來，集思廣益，思考一下。第二回合【張貼】標題：「最佳詮釋」，本文：哪一位同學詮釋得最好(自己除外)，其理由(20 個字以上)。第二回合【張貼】截止時間就是本議題在平台上的關閉時間。

2. 師生非同步教學討論議題：隨機抽樣方法的設計

第一回合請於 D+3 日晚上 12 點以前從「議題討論」區【張貼】標題：「隨機抽樣」，本文：請具體舉例說明：調查對象、調查地點、調查時間和如何執行隨機抽樣(具體詳細程序)(50 個字以上詮釋)。

期望可以透過議題討論資訊的交流，相互激勵，提升學習效益。待有 50 篇第一回合【張貼】回應後或第一回合【張貼】時間結束後，檢視其他同學的回應內容。自己靜下心來，集思廣益，思考一下。第二回合【張貼】標題：「最佳詮釋」，本文：哪一位同學詮釋得最好(自己除外)，其理由(20 個字以上)。透過同學之間的討論分享，可以具體提升學習效益。加油。第二回合【張貼】截止時間就是本議題在平台上的關閉時間。

3. 非同步教學中學習者之間討論議題：數值型態

第一回合請於 D+3 日晚上 12 點以前從「議題討論」區【張貼】標題：「四種數值型態」，本文：選定一個特定調查議題蒐集資料(需要文字說明)，針對四種數值型態依序設計四個問題與其對應的回答選項(方式)，分別符合四種數值型態。並且詳細說明如何判別該問題與其對應的回答選項(方式)屬於哪一種數值型態(20 個字以上)。書中範例不得使用，前面同學回應的內容【問題與回答方式(選項)不得使用】。

期望可以透過議題討論資訊的交流，相互激勵，提升學習效益。待有 40 篇第一回合【張貼】回應後或第一回合【張貼】時間結束後，檢視其他同學的回應內容。自己靜下心來，集思廣益，思考一下。第二回合【張貼】標題：「最佳詮釋」，本文：哪一位同學詮釋得最好(自己除外)，其理由(10 個字以上)。透過同學之間的討論分享，可以具體提升學習效益。加油。第二回合【張貼】截止時間就是本議題在平台上的關閉時間。

重點整理

統計學(Statistics)

在社會科學和自然科學中，非常重要與實用的一門科學，主要是運用於蒐集可靠和客觀的資料、彙整資料、組織資料、分析資料與解釋資料所蘊含的意涵，並嘗試運用數值去推論和瞭解真相與事實。

母體、族群、群體(Population or universe)或母群體(Parent population or parent universe)

在特定研究議題欲研究調查的可能對象，由一群具有某種共同特徵的基本單位所構成的群體(組合)。母體中基本單位的數量一般利用 N 符號代表。

基本單位(elementary units)、研究個體、個體(Individual)、元素(Element)

係指母體中可以被觀察或調查之不可分割、個別份子或單獨(單一獨立)的個別成份。

樣本(Sample)

母體的其中一部份基本單位所組成。樣本中基本單位的數量一般利用 n 符號代表。

研究變數或變數(Variable)

在母體或樣本中每一個基本單位，具有多種對研究觀察或調查有價值(興趣)的特徵或特質(characteristics)，稱為變數，經常使用大寫字母代表，例如：變數 X ，該變數的數值稱為變量(x)，使用小寫字母代表。

觀察值或觀測值(Observation)

對特定基本單位之特定變數進行測量或觀察，而獲得的測量值。

母體參數、母數、母體參數(population parameter)、介量或參數(Parameter)

代表母體中特定特徵(characteristic)、屬性(attribute)或變數的數值(測量值)。

統計值、統計量(statistics)、樣本統計值、樣本統計量(sample statistics)或估計值(Estimate)

由樣本獲得的特定特徵、屬性或變數的數值(測量值)，用以推估母體參數。

名目尺度、名義尺度(Nominal scale)或類別尺度(Categorical scale)

名目尺度是四個量表水準中，**最基本**的一種類型。此類型的問題，讓受訪者回答時，只需要提供一些名稱(names)、符號(label)、類型(category)等型態的選項。

順序尺度、次序尺度、序位尺度或等級尺度(Ordinal scale)

順序尺度允許受訪者依據回答選項之間，可以進行相對大小的比較，選擇一個與其感受相同程度選項。

等距尺度、區間尺度或間距尺度(Interval scale)

具有類別和順序尺度特性，並可比較差距。前後距離相等，間隔(Interval)可以用來加減，但沒有倍數關係，所以沒有絕對的 0 點、有相對的 0 點、0 點、原始點或基準點可設定於任一點(數值)。

等比尺度、比例尺度、比率尺度、比值尺度(Ratio scale)

具有類別、等級、等距尺度特性，有絕對的零點，且可以做倍數計算，具有可乘性(multiplicativity)。

質化資料、質的資料、質性資料、定性資料或屬性資料(Qualitative data)

調查變數的特徵屬於非數值型態。需利用一組的標記、類別、性質或名稱以區別每個基本單位的特徵(characteristic)、屬性(attribute)。

量化資料、數量資料、量的資料、量性資料、定量資料或屬量資料(Quantitative data)

可以使用數值表示的統計資料，如銷售量、銷售金額、收入、支出、身高、體重、年齡、小孩數量等。

橫向面資料、橫斷面資料(cross-sectional data)

同一時間或在同一特定時間區間內蒐集到的資料。一般指的是在此特定時間點，欲研究觀察之變數的觀察值。

時間序列資料、時間數列資料(time series data)

在一段時間區間內，連續對特定變數的觀察值。主要是希望呈現不同時間點，該特定變數的觀察值之變化情況。

縱橫資料(panel data)

同時包含橫向面資料和時間序列資料。

試驗設計、實驗設計(experimental design; design of experiments)

運用試驗過程(資料蒐集)中之重複性與隨機性，使研究變數(實驗變數、自變數)以外之其他變數(含已知與未知)對觀察變數或觀測變數(依變數)的影響相互抵銷，以單純化評量研究變數(實驗變數、自變數)對觀測變數(依變數)的影響效益，以提升分析結果的準確性。

敘述性統計(學)、敘述統計(學)(Descriptive statistics)、描述統計學

對於母體或樣本中特定變數的大量資料(觀測值)，可透過數值、統計圖或統計表將大量的資料，清楚和直接的呈現資料所代表的訊息或意涵。

推論性統計(學)、推論統計(學)(Inferential statistics)、統計推論(學)(Statistical inference)、歸納統計學(Inductive statistics)

在絕大部分的情況下，科學性研究皆受制於人力、物力、財力、時間等資源的限制，無法對母體中所有的基本單位進行評量或調查(普查)，只能夠透過抽選母體中的部分基本單位進行評量或調查。

數理統計(學)、理論統計(學)、理論統計

運用數學相關理論闡述統計方法的原理，推導各種統計公式，研究各種統計推論的方法與程序為整個統計學的理论基礎。

應用統計學(Applicative statistics)

運用統計相關理論和統計方法在各種科學領域、學術及管理上，以解決實際面臨的統計問題者。

母數統計學、有母數統計學(parametric statistics)、母數統計法(parametric method)

母體欲觀察研究之變數屬性或特徵分布屬於常態分布或趨近於常態情況，分析研究的樣本資料時，在統計的假設條件下，事先建立分析模式，此分析模式中皆會有未知的參數(parameter)。透過樣本的資料，進行參數(parameter)的估計與檢定確定分析模式中的參數數值，利用此模式進行分析或預測。

無母數統計學(nonparametric statistics; non-parametric statistics)、無母數統計法(nonparametric method)

適合使用於母體欲觀察研究之特定變數的屬性或特徵分布不為常態也不易轉換為常態、母體分布情況不明(不確定)、樣本數量 n 小、類別或順序變項。運用樣本資料中的大小、順序或等級之特性，進行統計推論之方法。