

單元：資料與測量的尺度

統計是門有系統的學科，大致包含三個部分，一是如何收集資料，二是如何整理資料，三是如何分析資料，底下我們先來介紹何謂『資料』。

1-1 資料

資料(data)，也可稱為數據，它是被收集分析以及彙整的一些事實或數字。對於某一特別研究所收集到的資料，我們稱為**資料集**(data set)。對於所收集到的資料之個體稱為元素(elements)，如一個公司，一個人等皆可當做元素。元素中我們有興趣的某一個特性稱為變數(variable)，如性別、學歷等。因資料的收集是根據某一個元素中每一個變數的測量而得到的，所以我們收集到某一個元素各變數之測量值則稱為觀察值(observation)，底下我們以10個人同時喝同一品牌煮出來的咖啡，並寫下一份咖啡問卷，將問卷結果彙整如下：

實驗者	性別	年齡	調味品種類		咖啡的味道	咖啡的濃度
			糖	奶精		
甲	男	25	有	有	剛好	剛好
乙	男	34	無	有	稍苦	剛好
丙	女	27	有	無	稍苦	稍濃
丁	男	30	有	有	稍酸	剛好
戊	女	24	有	有	剛好	稍濃
己	女	39	有	有	剛好	剛好
庚	女	34	無	無	太苦	稍濃
辛	男	19	無	有	稍苦	剛好
壬	女	20	有	有	剛好	剛好
癸	男	40	有	無	稍苦	剛好

則此結果便構成一個資料集，而資料中每一個實驗者都是一個元素，所以此資料集共有10個元素。資料集內有6個變數，分別是性別、年齡、調味品種類(糖)、調味品種類(奶精)、咖啡的味道、咖啡的濃度。當我們收集到第一個元素(甲)，則其觀測值依序為男, 25, 有, 有, 剛好, 剛好。要注意的是，有一個變數(年齡)的資料是數值，而其他變數的資料都是非數值的。換言之，我們收集到的資料並非完

全都是數值的，也可能是非數值的。

1-2 測量的尺度

任何測量必須有測量的準則和依據，例如量一張矩形桌子的長度，可以用原子筆長來量，或可用一般公制的尺來量，總要有個依據才行。這個做為測量的準則或依據，我們稱之為『尺度』。在測量時，如採用不同的尺度，則對變數的描述或說明提供不同的信息，以上述量矩形桌子長度為例，採用不同尺度測量的結果，可能是6倍的原子筆長，也可能是90公分。雖然桌子的長度不變，但因測量尺度不同，故對桌子長度的說明也不一樣。一般而言，變數的測量尺度有下列四種：

(1) 名義尺度, (2) 順序尺度, (3) 等距尺度, (4) 等比尺度。分別介紹如下

(1) 名義尺度(nominal scale)

名義尺度是根據變數之元素特徵與屬性不同，給予名字或代號，做為一種標記，進而將特徵或屬性相同者歸為類別，故也稱為類別尺度(categorical scale)。如以下變數均為類別尺度：

- a. 人之性別：可區分為『男性』、『女性』；
- b. 婚姻狀況：可區分為『未婚』、『已婚』、『喪偶』、『離婚』；
- c. 宗教信仰：可區分為『佛教』、『道教』、『基督教』、『其他』；
- d. 休閒嗜好：可區分為『戶外休閒』、『室內休閒』；
- e. 就業情況：可區分為『有』、『失業』；
- f. 家中電話區域號碼：可區分為『02』、『03』、『06』、『07』、…；
- g. 咖啡的味道：可區分為『太苦』、『稍苦』、『剛好』、『稍酸』、『太酸』；

我們也可以用數字來代表類別尺度非數值的情況，如針對婚姻狀況而言，可用數字1代未婚，2代表已婚，3代表喪偶，4代表離婚。這裡的數字1,2,3,4只是一種辨別之用。需特別注意，類別資料(變數為類別尺度的資料)不適用加、減、乘、除等符號運算，即使類別資料為數字，做算術運算是沒有意義的。

(2) 順序尺度(ordinal scale)

當資料有類別尺度的性質且這些資料的排序是有意義的。則可採用順序尺度。如某餐廳會在每一餐桌上放一份如下之問卷：

親愛的貴賓：

承蒙您的光臨，為了提供更好的品質，盼您提供寶貴的意見，敬請不吝批評，非常謝謝您。

請對餐廳以下項目予以評分：

環境	優良	好	普通	差	料理	優良	好	普通	差
清潔	☐	☐	☐	☐	味道	☐	☐	☐	☐
舒適	☐	☐	☐	☐	菜色與擺盤	☐	☐	☐	☐
氣氛	☐	☐	☐	☐	食物種類	☐	☐	☐	☐
服務	優良	好	普通	差	其他意見：				
親切 禮貌	☐	☐	☐	☐					
效率	☐	☐	☐	☐					
專業知識	☐	☐	☐	☐					

用餐日期：____年____月____日 用餐時間：☐ 早餐 ☐ 午餐 ☐ 下午茶 ☐ 晚餐

姓名：_____ 性別：☐ 男 ☐ 女

年齡：☐ 24 以下 ☐ 25-34 ☐ 45-54 ☐ 55-64 ☐ 65 以上

住家地址：_____

公司：_____

電話：(H)_____ (O)_____

填妥後，請將此問卷調查表交給餐廳經理，謝謝！

在此問卷中，針對料理部分為例，每位顧客分別根據料理中三個不同的變數(味道、菜色與擺盤、食物種類)給予一適當的等級，對每一個變數都『優良』、『好』、『普通』、『差』四種等級，而每一變數都須從四種等級中勾選一種。當該餐廳經理收集到每位顧客所提供的資料(意見)後，可根據『菜色與擺盤』按『優良』、『好』、『普通』、『差』給予分類或排序。同理，該餐廳經理也可根據『用餐時間』按『早餐』、『午餐』、『下午茶』、『晚餐』給予分類或排序。

順序尺度如同類別尺度，可能是數值的或非數值的，對於數值的，並不適用加減乘除四則運算；若是非數值的，依然可以用數字來代表，如餐廳問卷中，可用數字1代表優良、數字2代表好、數字3代表普通、數字4代表差。

(3) 等距尺度(interval scale)

當資料有順序尺度的性質，且可以比較二個元素之間的差異。則可採用等距尺度。如觀察某日三個地方的最高溫度，其結果如下：A地攝氏15度，B地攝氏28度，C地攝氏34度，若用此溫度來排序，可以發現C地最溫暖(溫度最高)，A地最冷(溫度最低)，除此之外，兩個資料值之差是有意義的，如C地比A地高

$34-15=19$ 度, B地比A地高 $28-15=13$ 度, C地比B地高 $34-28=6$ 度。又如考試的成績也是等距資料的一種。注意的是, 等距尺度的資料永遠是數值的。

(4) 等比尺度(ratio scale)

當資料有等距尺度的所有性質, 且兩個資料值的比是有意義的, 則可採用等比尺度, 如考慮買A牌的汽車需花費150萬元, 買B牌的汽車需花費75萬元, 若以0表示車子免費, 則從這兩輛車的價格來看, A牌的汽車價格是B牌的汽車價格的兩倍。又如, 人之身高, 可以採用等比尺度來測量, 以0代表沒有身高, 0以上的不同數值代表實際的身高, 而身高200公分即為身高100公分的兩倍。體重或年齡的測量也是如此。一般說來, 只有物理特徵的測量(如長度、重量等)比較可能採等比尺度。注意的是, 等比尺度的資料都是數值資料, 且要明確指出『零點』代表的意義。

事實上, 我們可以將依測量尺度之不同, 將資料區分成兩類: 第一類為屬性資料(qualitative data)依照資料的類別給予名字, 屬性資料的類別分辨可用非數值的方式, 或是數值的代號, 通常是由類別尺度與順序尺度中得到, 第二類為計量資料(quantitative data)是指我們有興趣的量之多寡, 通常由等距尺度或等比尺度中獲得。

生活中統計的實例

1. 某公司提供會員可郵購音樂CD, 該公司針對新加入會員做了一份問卷調查, 其中有些題目為:
 - a. 您的年齡為何?
 - b. 你的學歷為何?(有4種, 包括國中以下、高中或高職、大專或大學、研究所以上)
 - c. 在您家中共有多少位成員(包含您本人)
 - d. 您的興趣是哪一種音樂(有10種, 包括熱門、輕音樂、爵士、...)
 - e. 在過去六個月您買了多少片音樂CD

試指出上面問題的答案, 適合採用那一種測量尺度? 又有那些屬於屬性資料? 那些是計量資料?

[解]:

- a. 可採用等比尺度: 因年齡為一數值, 且可以做排序, 又兩個人之年齡之差為有意義, 且兩人年齡之比也有意義。
- b. 可採用順序尺度: 因學歷只能從這四種中選一種, 且也可以按學歷高低排序。

- c. 可採用等比尺度：因家中成員人數為一個數值，且可以人員多寡排序，又兩個會員之家中成員之差或之比都是有意義的。
 - d. 可採用類別尺度：因從10種音樂之中，任選幾種。可依喜歡音樂不同加以分類。
 - e. 可採用等比尺度：因購買CD的片數為一數值，且可依CD片的多寡給予排序，又兩個會員所擁有的CD片之差或之比都是有意義的。
- 屬性資料有b,d兩個問題的答案，計量資料有a,c,e三個問題的答案。

2. 某公司出版資訊圖書，書的最後一頁有一張讀者服務卡，形式如下：

讀者服務卡	
謝謝您購買這本資訊圖書！為加強愛用者提供更好的服務，請詳填本卡各欄，寄回本公司(免貼郵票)，本公司將依您的寶貴建議，不斷地追求進步。	
書號：ABC-1234	
一、您的年齡：_____	
二、您的職業： ☐ 軍 ☐ 公 ☐ 教 ☐ 工 ☐ 商 ☐ 學生 ☐ 其他	
三、您的最高學歷： ☐ 研究所 ☐ 大學 ☐ 專科 ☐ 高中職 ☐ 國中以下	
四、您從何處得知本書： ☐ 書局 ☐ 電腦商場 ☐ 學校 ☐ 雜誌 ☐ 其他	
五、您通常以何種方式購書： ☐ 書局 ☐ 電腦商場 ☐ 學校集體採購 ☐ 網路購書 ☐ 郵購 ☐ 展覽 ☐ 其他	
六、您對本書的滿意程度：	
1. 內容表達：	☐ 很滿意 ☐ 滿意 ☐ 尚可 ☐ 不滿意 ☐ 很不滿意
2. 編輯校對：	☐ 很滿意 ☐ 滿意 ☐ 尚可 ☐ 不滿意 ☐ 很不滿意
3. 紙張印刷：	☐ 很滿意 ☐ 滿意 ☐ 尚可 ☐ 不滿意 ☐ 很不滿意
4. 封面設計：	☐ 很滿意 ☐ 滿意 ☐ 尚可 ☐ 不滿意 ☐ 很不滿意
七、其他建議：	

試根據這張讀者服務卡，判斷前六大項問題，適合採用那一種測量尺度？那些是屬性資料，而那些是計量資料？

[解]

第一大項：等比尺度，第二大項：類別尺度，第三大項：順序尺度，
 第四大項：類別尺度，第五大項：類別尺度，第六大項：順序尺度。
 屬性資料為：第二、三、四、五、六大項問題

計量資料為：第一大項問題

3. 某銀行為了感謝該行信用卡的卡戶常久以來的支持與愛護，特別舉辦多項卡友優惠活動：

優惠一：新辦理開戶，並使用該帳戶扣繳信用卡款，可獲贈『咖啡對杯一組』

優惠二：新申購基金達新台幣5萬元以上，可獲贈『咖啡對杯一組』

優惠三：新開戶+申辦信用卡自動扣繳+申購定期定額基金，可獲贈『精緻咖啡機一台』

該銀行並將此優惠活動訊息寄給該行的卡友，並隨信附上一張如下之問卷：

親愛的卡友客戶，您好！					
感謝您多年來對本行的支持與愛護，本行特別舉辦多項卡友優惠活動(詳情請見信中優惠活動訊息介紹)。此外，由於您的寶貴意見，將是我們進步的來源，請您撥冗填寫此份問卷，填妥後將此問卷交予本行各分行的服務人員，謝謝！					
1. 請您對下述敘述是否同意					
	非常同意	同意	沒意見	不同意	非常不同意
a. 我很滿意卡友優惠活動的設計	☐	☐	☐	☐	☐
b. 我很滿意服務人員態度	☐	☐	☐	☐	☐
c. 我會將貴行當做主要往來的銀行	☐	☐	☐	☐	☐
d. 我會想知道更多金融商品資訊	☐	☐	☐	☐	☐
2. 請問您目前共有多少張信用卡(包含他行)：_____張					
3. 您覺得本行卡友優惠活動比其他銀行的卡友優惠活動： ☐ 好 ☐ 差不多 ☐ 差					
4. 請問您通常會以何種方式查詢您的信用卡相關業務：					
☐ 電話 ☐ 網路 ☐ 銀行 ☐ 未曾查詢過					
5. 請問您希望我們以何種方式提供您各項金融商品的資訊：					
☐ E-mail ☐ 手機簡訊 ☐ 信件 ☐ 信用卡對帳單 ☐ 不希望收到訊息					
姓名：_____ 身分證字號：_____					
E-mail：_____ 手機號碼：_____					

試根據這張問卷，判斷這五個問題，那些是屬性資料，那些是計量資料。

[解]：屬性資料：1,2,4,5；計量資料：2。

4. 某公司想了解員工是否能分辨咖啡的好壞，而做了一份問卷，該公司準備了四種咖啡豆，每天各煮一種，一週煮了四天，共煮了3週，並控制每次煮的咖啡豆與水的重量。每位品嚐咖啡的員工，都需填寫如下問卷：

咖啡實驗問卷

1. 今天您品嚐咖啡的時間：____月____日下午____時
2. 性別：☐ 男 ☐ 女
3. 年齡：☐ 19 以下 ☐ 20-29 ☐ 30-39 ☐ 40-49 ☐ 50-59 ☐ 60 以上
4. 請問你平均一週喝多少杯咖啡：☐ 0-1 杯 ☐ 2-3 杯 ☐ 3-4 杯 ☐ 5 杯以上
5. 請問您覺得今天咖啡味道：☐ 太苦 ☐ 稍苦 ☐ 剛好 ☐ 稍酸 ☐ 太酸
6. 請問您覺得今天咖啡的濃度：☐ 太濃 ☐ 稍濃 ☐ 剛好 ☐ 稍淡 ☐ 太淡
7. 請問您今天喝的咖啡有無加調味品：☐ 無 ☐ 有，加了 ☐ 糖 ☐ 奶精
8. 整體而言，請問您覺得今天咖啡的等級：
☐ 非常差 ☐ 差 ☐ 普通 ☐ 不錯 ☐ 非常不錯

試決定這份咖啡實驗問卷中，各問題那些採用類別尺度較適合，而那些採用順序尺度較適合。

[解]：類別尺度：性別、有無加調味品

順序尺度：時間、年齡、平均一週喝咖啡的杯數、咖啡的味道、咖啡的濃度、咖啡的等級。