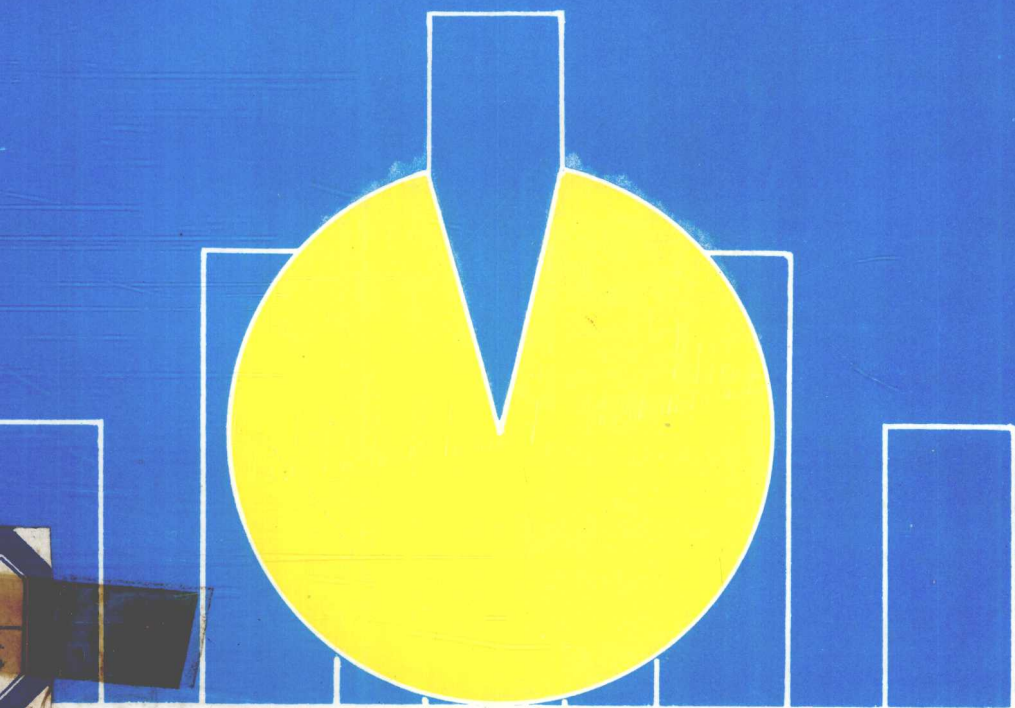


高二機率統計學最佳讀本 專科以上統計學教科用書

敘述統計學

呂永聰 編著

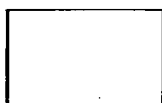


協進圖書有限公司印行

敘述統計學

民國七十五年五月初版

版權所有



翻印必究

編著者：呂永聰

發行人：柯水源

發行者：協進圖書有限公司

台北市羅斯福路三段316巷9弄4號

電話(02)393-8837

郵政劃撥：帳戶 協進圖書有限公司

帳號 0019089-0

行政院新聞局登記證局版台業字第0045號

定價新台幣110元

序 言

現代是一個科學昌明的時代，隨著科學的進步，統計學的應用也日趨廣泛，從日常生活以至於人文科學與自然科學的研究，有不少與統計學發生關聯，例如報章、雜誌，和廣播裏有統計報導，公司、銀行，或政府機構中也有很多資料要用統計方法來整理和分析，某些商業、經濟、財政，或教育的決策也都運用到統計學先行研究，在生物學、醫學，和氣象學上統計學理的運用亦很普遍，故身為現代人實有必要多學些統計學的知識。

敘述統計學是統計學中最基本且最淺顯的知識，也是統計學的入門，其所涉及之數學學理並不深，宜於專科或中學的學生來學習，某些較簡單的統計內容也可以為小學生所吸收。

本書雖只是一小冊，除能扼要地將敘述統計學的內容與精神包攬外，還融匯了最新的統計知識與方法，如莖葉圖表與方格圖等，我相信讀者對這些新的知識與方法一定會感到格外的興趣。

本書的原稿曾經國內外三十多位專家學者批閱，提供意見，作為修改之參考，並在國內商專二年級試教三次，根據學生的學習成果及其反應，修訂完成，相信本書一定很適於專科或中小學的學生來閱讀與研究，也可作為大學生和教師的參考資料。

本書倘仍有疏漏處，還請多多指正，不勝感激。

呂 永 聰 謹識于台北商專

民國七十五年四月

目 錄

第一章 緒 論

§ 1-1 統計學的分類	1
§ 1-2 資料層次的分類	3
§ 1-3 常數與變數及變數的分類	8
本章提要	9
問題討論	10
統計學簡史	11

第二章 資料的搜集、整理、陳列與顯示

§ 2-1 群體與樣本	15
§ 2-2 次數分配與累積次數分配	17
§ 2-3 相對次數分配與累積相對次數分配	21
重點提要	34
習題 2-1	37
統計趣談	41
§ 2-4 統計圖——次數分配的圖示法	41
2-4.1 折線圖與次數多邊形圖或相對次數多邊形圖	42
2-4.2 累積次數分配圖與累積相對次數分配圖	50
2-4.3 條形圖與直方圖	52
2-4.4 圓餅圖	54
重點提要	55
習題 2-2	56

統計疑義	59
§ 2-5 莖葉圖表	59
重點提要	76
習題 2-3	78
莖葉漫談	81

第三章 集中與離散的測度

§ 3-1 集中量數的類別	83
3-1.1 集中量數的類別	84
3-1.2 衆數	85
3-1.3 中位數	91
3-1.4 四分位數和百分位數	95
3-1.5 加權算術平均數	100
3-1.6 修整平均數	101
重點提要	102
習題 3-1	106
日常生活統計用語	110
§ 3-2 離散量數	111
3-2.1 資料分佈的離差	111
(1)全距	111
(2)均互差	111
3-2.2 以中位數爲集中量數時	112
(1)四分位差	112
(2)離中差	113
3-2.3 以算術平均數爲集中量數時	113
(1)離均差	113
(2)標準差	114

3-2.4 相對離散量數	117
重點提要	119
習題 3-2	123
習題解答	125

第一章 緒 論

統計學的英文字是 **statistics**，它的起源可以追溯至一、兩百年前，然而統計學理的蓬勃發展與運用的廣泛可以說是近半世紀以來的事，不論是自然科學或是人文科學，有不少的地方都以統計學理作為研究問題的工具，在自然科學方面如氣象的預報、藥物療效的研究，在人文科學方面如各種抽樣調查與測驗結果的分析，無不利用統計方法去解決問題，美國著名的蓋洛普民意調查也是使用了統計學上抽樣、推測的方法對選民的投票意向所作的分析與估測，其預測的結果往往和未來的事實相差無幾。

在工商業上統計方法也運用的相當廣泛，例如工業產品的品質管制、商品的檢驗、商情的預測、廣告效度的調查等，在經濟上如匯率的調整、國民所得的預測等，在保險公司裏有關保險費的計算、死亡表的制定等，通通和統計學息息相關。利用統計方法作**定量**分析所得的結果，可以幫助工商領導人對未來不定的情況作**判斷與決策**。其他較簡單的事情如資料之整理、摘要，與報告也會用到統計學上歸類、劃記，與圖示的方法。所以統計方法在工商業上是相當重要的。

即使在日常生活中，我們在報章雜誌或電視收音機裏，也會看到或聽到一些和統計有關的報導，因之生活在現代社會中的每個人都有必要具備一些統計的知識。

1—1 統計學的分類

統計學通常可分成**敘述統計學**(Descriptive Statistics) 和**推測統計學**(Inferential Statistics)。茲分述如下：

2 敘述統計學

(一)敘述統計學

在搜集了資料 (Data) 以後，爲了使它們系統化且能簡明扼要的把它們陳列出來，我們可以使用統計學上歸類，與劃記等方法，將原本雜亂無章的資料整理出來，並用圖表顯示之，使人一目瞭然，對這群資料有個印象，並從而去發現它們的共同點及其分佈的特性。對於數值的資料 (Numerical Data) 我們還可求其各種的平均數 (Mean)、中位數 (Median)，和眾數 (Mode)，以得知資料集中的情形。以上這些工作都是屬於敘述統計學的範圍，也是統計上最初等的工作，敘述統計學亦稱描述統計學。

例如某大專人事室爲了使人瞭解校內教師的職位狀況，就用圖示法來標明教授、副教授，及講師的人數，並區分成男女之別，使人一目瞭解校內男女教師中各職位的人數有多少。此外也可以用圖表標示教師年齡、身高等的分佈情形，同時亦可求其平均年齡，和平均身高，作爲人事上的參考資料。以上的這些工作在其他的機關或公司裏，也是常見的。

(二)推測統計學：

適用於推知甚廣大的一整體的特質或現象以及未發生之事物可能發生的情況。

有時候我們爲了研究一整體 (Population，或譯成羣體、母體、母羣，或全體) 的某些特性或現象，由於群體所包含的個體太多或太廣，爲時間、財力、精力、能力，或其他客觀因素所限，我們無法將群體中各個個體的資料拿來整理、分析，所以就用抽樣 (Sampling) 的方法，從群體中來選取有代表性的個體或元素，作爲研究的樣本 (Sample)，由樣本所得到的資料來推知群體的現象或特性，例如燈泡製造廠爲了瞭解所產燈泡的品質與壽命，當然不可能將所有的產品都拿來試驗，所以只得抽取某些燈泡來作研究，從這些燈泡樣品的品質與壽命來推測其他所有燈泡產品的品質與壽命，這就要運用到統計學上機率與推測的學理。

另外對於未來不定現象的推測，也要用到統計方法，例如某公司為了預測某產品的銷路情況，以作為決定產量的參考，就必須作市場調查，從現有的資料來推知未來可能的銷路狀況。

以上只是一個概略的分法，好讓初學者略知統計學的概況而已，在商業上統計學有一個極重要的目的就是提供決策 (Decision making) 的參考，也就是利用統計方法整理、歸類、分析、歸納，或推測的結果，來決定進行的方案，以得到最高的利益。在學術上，很多商科學門，也以統計學作為研究的工具，描述統計學在處理商務方面的用途亦甚廣，所以商業或工業學校的學生不可不學統計學，而有志從事工商業的人士也不可不懂統計學上最基本的知識。本書就是要介紹統計學上最基礎的知識與方法，使學習的人獲得統計學上一些初等的且基本的概念，並能運用統計方法來處理某些工商問題，學習的人務須將這部份融會貫通，不但有助於將來統計學的學習，而且對修習工商科目也有裨益。

由於統計學上所處理的資料大多是數字，所以統計學就與數學發生了關聯，而統計學也就奠基在數學學理之上，來對量化的資料，作定量分析。在本書裏我們所談的統計資料大多指數值資料而言。

1—2 資料層次(Data Levels)的分類

為了應用統計學作研究的工具，研究者就必須收集相關的資料，再加以整理、分析。史蒂文斯 (S.S. Stevens, 1951) 按照各種數值資料的性質，將資料分成四個類等。有些數值資料只是人或物的代號，兩數目之間，非但沒有比率、大小的關係，而且沒有次序關係，這種的數值資料就稱為名義性的 (Nominal) 或分類性的 (Categorized) 資料。較佳一點的資料就是順序性的 (Ordinal) 資料，這類的數據有比較大小、優劣，或先後的關係。更佳一點的資料就是等距性的 (Interval) 資料，或稱等差性的資料，這類的資料，可由

4 敘述統計學

兩數據得知其差異有多大。最佳的資料，不但可以顯示出兩數據差異的大小，而且代表著實體或實質的比率關係，那就是**比率性的**(Ratio)資料。茲分別說明如下：

(1)名義性的(Nominal)資料

這類的數據只不過是事物的代號而已，有的只是用來命名或分類罷了，兩數據間並無比較之關係。例如球員的背號是任取的代號，並非按球員的球技、身高，或體重等而編取的號碼，其所代表者只是該球員而已，而沒有其他的意義。

某市場調查員爲了調查藍寶、汰漬、白蘭、世霸等四種洗衣粉的銷路情況，爲了方便起見，分別給了這四種洗衣粉予 1, 2, 3, 4 等四個編號，則這四個數據 1, 2, 3, 4 亦僅是不同牌洗衣粉的代號而已，並不代表其他的意義或兩種洗衣粉間的任何關係，故僅是名義性的資料而已，既無比率性和等差性，亦無順序性。



(2)有順序性的(Ordinal)資料

這類的資料可以分出所代表之人或物在某性質 (Characteristic) 上的**優劣、高低**，或**先後**之關係。例如按學業成績來排定班上學生的名次，這第一名、第二名、第三名……等等的一、二、三……等數目就代表著該學生在學業成績這方面的高低，當然第一名的學業成績比第二名的高，第二名的學業成績又比第三名的高，但我們不能單由名次而知道相隔兩名間的差異量有多大，更不能知道第一、二名

間的差異是否與第二、三名間的差異相同，或有何大小關係。

某電腦經銷商銷售四種微電腦產品，在去年十月間的銷售量自高而低，依次為 IBM、蘋果牌、小神通，和小教授，今分別用 1，2，3，4 四個數字來顯示它們在去年十月間銷售量的高低，則這些 1，2，3，4 等四個數據就是有順序性的數據，由於 2 這個數字代表蘋果牌微電腦的銷售量排第二，而 4 這個數字代表小教授微電腦的銷售量排第四，所以我們可以得知在去年十月，該經銷商賣出蘋果牌微電腦的量比賣出小教授微電腦的量還多，但是如果我們要知道各廠牌的微電腦在去年十月間該經銷商銷售量之多寡及大小差異等情況，就不能僅由我們剛才所用的 1，2，3，4 四個數據而得知了，這也就是說我們不但要知道和事物有關的數據或資料，還要瞭解那些數據或資料所指的是什麼性質（Characteristic）。

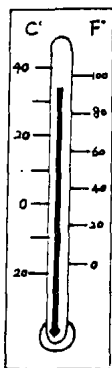
(3)等距性的(Interval)資料

這類的資料不但有大小、高低，或優劣等順序關係，而且由兩數據的差可以知道差距有多大，這類的資料通常具有一個**共同的基本單位**，而且此基本單位的量是一定的且是已知的，例如科學上規定溫度的基本單位——攝氏或華氏一度——是人人通曉而且其量是固定的，所以如果我們獲得了一群有關溫度的資料，我們不但可以由溫度的數值來比較兩溫度的高低，也可由兩數據的差來知道兩溫度的差異量有多大，且差異量間亦可比較大小。例如攝氏二度的溫度當然比攝氏十度的溫度低，且低了八度，這和 34°C 比 42°C 低了 8°C 是一樣的，亦即在溫度計上的差距是相等的，因之有關溫度的資料皆是等距性的資料，可惜它們並沒有比率性，因為我們不能說 10°C 的溫度是 2°C 的溫度之 5 倍。

另外有關日期的資料也有等距性，例如從今年一月三日至一月八日共六天，這和同月九日至同月十四日所經過的時間是一樣長的。這些資料三日、八日、九日，和十四日都是等距性的資料。

6 敘述統計學

等距性的資料在工商業上也會遇到，除了上面所提到的溫度與日期之外，如金額和度量衡的數據皆有等距性，並且金額以及由度量衡所得到的數據亦有比率性，茲說明如下：



今天的溫度比
昨天高了 2°C

(4)有比率性的(Ratio)資料

這類的資料不但可由其數據知道兩者差別之大小，而且兩數之比率可以代表其所指事物**某性質的比率**。如上所述，由度量衡所得的數據皆是有比率性的資料，例如6公尺長一定是2公尺長的3倍，100 C.C. 的容量一定是50 C.C. 的2倍。大致上來說，這類的資料皆有一個「絕對零點」(Absolute Zero)，例如0公尺的0所代表的就是沒有長度，也就是測量長度的「起點」，又如0公斤就表示沒有重量，它是稱重量的「基點」。這和前面所說的代表溫度或日期的資料數據略有不同，因為 20°C 的溫度並不一定是 17°C 的 $20 / 17$ 倍， 10°F 的溫度並不是 -2°F 的 -5 倍，日期三日也不是八日的 $\frac{3}{8}$ 倍。

例1 下例是有關張亨利、陳傑克、林吉米、黃貝琪、劉瑪莉等五位同班同學的資料：

	張 亨 利	陳 傑 克	林 吉 米	黃 貝 琪	劉 瑪 莉
學 號	71010	71237	71125	71089	71346
身 高	169 公分	170 公分	165 公分	157 公分	160 公分
體 重	58 公斤	70 公斤	53 公斤	46 公斤	45 公斤
體 溫	38.5 ℃	38.1 ℃	37.9 ℃	38.9 ℃	39.0 ℃
生 日	57.7.2	57.5.8	56.11.12.	56.10.7.	57.6.6.
第一學期的 名 次	4	30	21	15	11
上次月考的 數學成績	89	76	53	81	64
本學期第二 次的作文成 績	乙	甲	乙	甲	丙
跑一百米所 需之時間	18 秒	16 秒	15 秒	19 秒	17 秒

試問各組資料是屬於那一層次？

解答：有關學號這組資料是屬於名義性的資料，每個數字只是某一位學生的代號而已。但是如果註冊組的人員按照同班同學之學號的數值大小，排成點名的先後次序，則有關學號的資料又有順序性。

有關第一學期的名次和本學期第二次的作文成績這兩組資料皆是有順序性的資料，它們各別表示了第一學期的總成績和本學期第二次的作文成績之優劣而已，但並不能表示出差異量，所以不是等距性的資料。

至於這兒有關體溫和生日的資料則是等距性的資料，它們不但可以顯示出差異有多大，其差異量可用 1 ℃ 或 1 日作衡量。

的單位，但這些數據間的比率並不能代表著溫度或日期的比率

。

其餘各組的資料身高、體重、上次月考的數學成績，和跑一百米所需之時間等皆是有比率性的資料，身高指的是長度，體重指的是重量，數學成績指的是得分點數，跑百米所需之時間指的就是時間，由數據的比率可以得知所欲測之某性質（如長度、重量、時間等）的比率。

1—3 常數(Constant)與變數(Variable)及變數的分類

當我們要應用統計方法去解決問題或作決策時，我們首先要針對問題的需要去搜集資料，如果有關某性質的資料，不因所要研究之個體或事物的不同而有異，則這個性質就是**常數**(Constant)，或稱作**定元**或**恆性**。否則就是**變數**(Variable)，或稱作**變元**或**動性**。

例如我們如果要研究台北商專二年級學生的身高，我們當然要搜集相關學生的身高資料，則這兒的「校別」和「年級」是常數，因為這兩個資料對所欲研究之學生都是一樣的，而「身高」却是變數，因為學生的身高不全相等。假使我們另外又想研究性別之差異，在獲得了身高之資料以外，又知道了每個身高資料是男生的，或女生的，則「性別」也是變元。

此外關於變數的分類，可有甚多種的分法，在本書中我們只想談到**間斷性的**(Discrete)與**連續性的**(Continuous)的變數。通常來說，若所獲得的資料是名義性的或順序性的，則大多是間斷性的資料，而其所指的性質就是間斷性的變數，也可說是**非連續性的**(Discontinuous)。例如學生的學號是名義性的資料，即使 71203 和 71204 是某兩學生的學號，但此兩號之間的任何數目如 71203.3, 71203.6 等等皆不可能是學生的學號，所以學號是間斷性的變數。然而等距性的資料和比率性的資料大多是連續性的，例如有關溫度的

資料，任何兩數據間的數目皆有可能是溫度的計數，因之溫度是連續性的變元。但有某些等距性或比率性的資料並非連續性的，如人數資料有比率性，但無連續性。

本章提要

一、統計學大致上可分成二部分

- (一)敘述統計學：用劃記與圖表的方法來整理與歸類資料，或求得一群資料的集中與分佈概況，都是描述統計學的範圍。
- (二)推測統計學：以描述統計學和機率論作基礎，從分析有限個樣本來推測無限的大群體的某些性質，或預測未來可能發生的情況。

二、統計工作的步驟

- (一)搜集資料，(二)整理資料，(三)分析資料，(四)推測未來。

三、資料層次的分類

- (一)名義性的資料：其作用在分類或命名而已，所以資料間只有相等關係，而無順序、差量，或比率的關係。
- (二)次序性的資料：其效用在分別出某性質的優劣或先後、高低、多寡等，有順序或大小之關係，但不能顯示差量或比率。
- (三)等距性的資料：可以顯示出差距量的大小，這類的資料有一個共同的度量單位，所以由數據的差量可以知道差別有多大。
- (四)比率性的資料：在獲取這類資料之時，不但有一個共同的度量單位，而且也有一個絕對零點，故這類的資料可以反映出所測度之性質的大小、差異量，和比率。

四、變數與常數

- (一)變數是指可能因個體或事物而變化資料的性質。
- (二)常數是指不因個體或事物之不同而有異的性質。

五、變數的分類

- (一)間斷性的資料或變數指的是兩資料數據間的數值有可能不是所有研究之個體或事物的資料者。
- (二)連續性的資料或變數指的是任兩資料數據間的任何數值皆有可能是研究群體中某個體或事物的資料者。

問題討論

1. 指出下列各類資料的層次：

- (1)國民身份證號碼
- (2)同一條街的門牌號碼
- (3)某公司員工的薪資
- (4)叢林中樹木的高度
- (5)百貨店中各種杯子的容積
- (6)臺北火車站南下火車的開車時刻
- (7)臺北市納稅人的年所得稅額
- (8)臺北市公共汽車的路線編號
- (9)臺北市各個學校的教師人數
- (10)臺北商專所有學生的年齡

2. 在上題中那些資料是間斷性的？那些是連續性的？

3. 指出下列各情況的變元：

(1)某研究員獲得了臺北市十家公司總經理的月薪如下：

10 萬，3 萬 7 千，5 萬 6 千，9 萬 3 千，6 萬 4 千，
4 萬，5 萬 2 千，7 萬 1 千，3 萬 3 千，2 萬 8 千。

(2)自民國 63 年至 71 年國中應屆畢業生的升學率如下：

53.38 % ,56.69 % ,52.43 % ,51.19 % ,50.86 % ,
52.66 % ,54.86 % ,57.53 % ,60.00 % .

- (3)某市場調查員爲了研究臺北市某些文具行在上個月銷售雷諾、金剛、秘書、利百代等四種牌原子筆的銷售量，所以搜集了甲、乙、丙等三家文具行有關的資料如下：

	雷 諾	金 剛	秘 書	利百代
甲	123	112	93	101
乙	96	78	145	84
丙	106	93	81	67

以上諸數據的單位爲支。

統計學簡史

一、遠古雛形時代

在四、五千年前，某些古文明國家就有了統計資料，以供行政之用，例如在 西元前3050年，埃及國王爲了金字塔的工程，曾調查全國人口和財富。而我國在 西元前2200年 左右，夏禹王分天下爲九州，記載山川人民的分佈。此後數千年，中西各國對官民、兵馬、財源、土地等政治情形皆略有記載，可以說是有了統計的雛形，只不過僅局限於統計資料的搜集與整理而已，並無特殊的統計方法，而且應用亦不普遍。