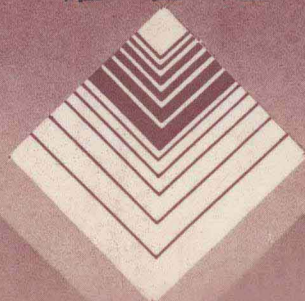


學術譯叢

商用統計學

詹世煌 編譯



五南圖書出版公司 印行

大 專 用 書

商 用 統 計 學

詹 世 煌 編 譯

五 車 圖 書 出 版 公 司 印 行

商 用 統 計 學

中華民國 72 年 2 月初版

編譯者 詹 世 煌
發行人 楊 榮 川

發行所 五南圖書出版公司
局版臺業字第 0598 號
臺北市銅山街 1~1 號
電 話：3916542 號
郵政劃撥：106895 號

基本定價： 6.67 元

印刷所 明文印刷廠

(本書如有缺頁或倒裝，本公司負責換新)

編譯者序

加拿大西安大略大學(University of Western Ontario)兩位教授，Tomas H. Wonnacott 博士和 Ronald J. Wonnacott 博士，合寫的「統計學導論」乙書，在美國甚為風行，各大學爭相採作教科書。本書就是根據這本統計學的原著，加以編譯改寫的，其中絕大部分保留著原著的內容與風貌，只是就其不適合國情的部分，加以改寫而已，希望讀者讀來更覺切當。

致學生

統計學是一門相當迷人的學問，它告訴你如何僅由幾扇窗之所見，來描述一個未知的世界。由此，你將獲得一些思考性的樂趣，而此一樂趣在此之前，你是從未有過的。

本書並非一本小說，因此不能以看小說的方式來讀它。雖然每位學生對教科書皆各有其自己的讀法，在此我們仍將說明由很多已有所成的學生所推薦的方式。首先仔細地讀一節，同時在最重要的部份下劃——線，但並不摘取困難的部份。對教科書中的一些例題，自己先試行求解，於經過苦思之後（且有希望得解），再參考書中的答案，而後繼續下一個問題。一旦做了此等問題的解答，它將拉你回到教科書中，而了解最重要的概念。由於問題解答通常能解釋以後的概念，因此學生應牢牢記住一項重要的法則：「做問題解答應永不落後」。本書中問題的計算已經盡可能的簡單，因此你可專注於概念上的了解。同時在使用實際資料時，我們亦試圖使問題能夠合理化。

在符號方面：有星號(*)者乃是較艱深的題目，因此可以省略。若該題目係重要的概念，則標以箭號(⇒)，若題目與前述的問題重

複，則以括號()表示，故此等題目可省略。

本書前半部扼要說明了各種概念間的關係，對此部份常加複習，不但可溫故，在未來的參考上亦深有助益。

致教師

由於本書的基本目的，乃是盡可能以簡單的方式來解釋統計概念，因此所需要的預備知識僅高中數學便已足夠。微積分僅在非使用它來論證不可的章節下方用之。但即使如此，這些章節亦經過設計，以便沒有修過微積分的學生亦能研讀。同樣地，本書不用抽象的理論，而用問題來例示、介紹新的素材；必要的理論僅在學生對已介紹之理論有一清晰、直覺的概念時，方加以說明。

本書另一個主要目的，乃是對教科書中常分開且孤立的章次，說明其各主題間的邏輯關係。具此類關係的例子如：區間推定和假設檢定； t 檢定和 F 檢定；變異數分析和用虛擬變數的迴歸等。對上述情形，我們有如下的二個動機：幫助學生了解基本的邏輯推理，和幫助他們求得實務問題的解答。

我們將迴歸模型置於高度優先，不僅是因為迴歸廣泛地為實用統計學家視為最有效的工具，最主要的，乃是由於迴歸係了解相關變異數分析和卡方等相關技巧的一個優良的「焦點」。自簡單迴歸推廣到複迴歸和非線性迴歸，我們亦予以相當的篇幅介紹。

雖然囿於初學者，本書的處理盡可能的簡明，但它仍包含了在其他初等統計學內通常所沒有的材料。例如，為對指數的內容有更深入的了解，我們相當簡要地介紹了經濟理論上的顯現性偏好。同樣地，時間數列一章不僅包含了一般的課題（如季節和趨勢的調整），而且亦對自我相關誤差和其如何處理有一簡單的討論。除此之外，我們亦介紹了譜相分析。此外有一章推定聯立方程式——對商學和經濟非常重要的課題。如同譜相分析，此章僅在較高等的教科書和論文中方說明。起先，我們認為這些課題不能在初等的水準時來介紹，但經過簡單化後，我們希望已能祛除它過去在大學部學生能力所及的學術上的神秘。

本書具有相當大的彈性。前面 15 章敘述基本的古典統計學；而後面 10 章為獨立的課題，包括上面所提過的主題，和無母數統計、卡方、抽樣設計、貝氏推定與賽局理論等。對此 10 章，教師們可依課程的需要斟酌章節來講授。本書為保持簡易起見，比較複雜的解釋和引證留在附註和標有星號的節次裏，其例題皆省略。為使更初等的學生能完全地跳過此處不讀而又不失連貫性起見，本書對此已有特殊的處理。

謝 言

最後，編譯者要感謝在各方面提供協助的人士，有了他們的指導與支持，本書才得以問世。在編譯的過程中，雖然已盡了最大的努力，但有瑕疵的地方，恐尚不能免，尚望統計學界的先進，能予以指正，俾在再版時有所改進。

編譯者 詹世煌

71 年 6 月 1 日
於政治大學統計研究所

商用統計學

目 錄

第1章 導 論	1
1-1 例 子	1
1-2 演釋法和歸納法	4
1-3 為何要抽樣？如何抽樣？	5
第2章 樣本敘述統計學	9
2-1 次數表和圖形	9
2-2 分配的集中趨勢	12
2-3 分配的離勢	16
2-4 線性變換	19
第3章 機 率	29
3-1 導 論	29
3-2 出象和出象的機率	31
3-3 事項和事項的機率	34
3-4 條件機率	42
3-5 獨立性	46
3-6 其他觀點的機率	49
第4章 機率分配	63
4-1 不連續隨機變數	63
4-2 均數和變異數	66
4-3 二項分配	70

2 目 錄

4-4	連續分配	74
4-5	常態分配	77
4-6	隨機變數的函數	83
4-7	附 註	85
第 5 章	兩個隨機變數	95
5-1	分 配	95
5-2	二隨機變數的函數	103
5-3	互變數和相關	105
5-4	二隨機變數的線性組合	108
第 6 章	抽 樣	123
6-1	隨機抽樣	123
6-2	樣本均數的動差	126
6-3	中央極限定理	129
6-4	抽樣之模擬(蒙地卡羅法)	132
6-5	0—1 變數	139
6-6	小母體抽樣	143
第 7 章	點估計	153
7-1	母體和樣本	153
7-2	推定量所應具有的性質	154
7-3	一致性	159
7-4	無母數推定簡介	161
第 8 章	區間推定	171
8-1	單一均數	171
8-2	σ^2 未知的 t 分配	177
8-3	二均數的差	181

8-4	比 率	187
8-5	單尾信賴區間	191
8-6	常態母體的變異數	192
第9章	假設檢定	205
9-1	利用信賴區間來檢定假設	205
9-2	機率值	208
9-3	古典的假設檢定	212
9-4	再論古典檢定	215
9-5	β 函數和檢定力函數	218
9-6	兩尾檢定	223
第10章	變異數分析	233
10-1	一因子變異數分析	233
10-2	信賴區間	242
10-3	二因子變異數分析	250
第11章	直線的配合	265
11-1	簡 論	265
11-2	配合一直線的評判標準	268
11-3	最小平方法	270
第12章	迴歸理論	279
12-1	數學模型	279
12-2	$\hat{\alpha}$ 、 $\hat{\beta}$ 的均數和變異數	282
12-3	高斯——馬可夫定理	285
12-4	$\hat{\beta}$ 的分配	286
12-5	β 的信賴區間和假設檢定	287
12-6	內插(區間推定)	290

4 目 錄

12-7	外插法	295
12-8	結 語	296
第13章	複迴歸	303
13-1	簡 論	303
13-2	數學模型	305
13-3	最小平方推定量	306
13-4	共線性	307
13-5	信賴區間和統計檢定	311
13-6	應有多少迴歸子	314
13-7	迴歸的解釋：其他事物不變下的解釋	316
13-8	簡單迴歸與複迴歸之比較	318
13-9	虛擬（0—1）變數	320
13-10	迴歸和變異數分析	324
第14章	相 關	341
14-1	簡相關	341
14-2	相關和迴歸	349
14-3	偏相關和複相關	359
第15章	非線性迴歸	369
15-1	變數非線性，但母數則否	370
15-2	母數為非線性，需要變換的情形	372
15-3	彈 性	377
第16章	無母數統計	391
16-1	符號檢定	392
16-2	來自順序統計量的信賴區間	394
16-3	二樣本的W檢定	397

16-4	隨機性的檢定	401
16-5	無母數檢定的優點	405
第17章	卡方檢定	417
17-1	卡方適合度檢定	417
17-2	列聯表	424
17-3	分配形狀之卡方檢定	427
第18章	最大概似推定 (MLE)	435
18-1	比率 π 之MLE	435
18-2	常態均數之MLE	439
18-3	常態迴歸係數之MLE	441
18-4	來自任意母體中任意母數的MLE	444
18-5	MLE及MME (動差推定法)	444
第19章	貝式推論	447
19-1	事後機率	447
19-2	母體比率 π 的事後分配	450
19-3	常態母體均數 U 的事後分配	455
19-4	貝氏迴歸	460
第20章	貝式決策理論	467
20-1	最適決策	467
20-2	點估計即——決策	470
20-3	做為決策的假設檢定	473
20-4	MLE 及貝氏推定	478
20-5	貝氏方法的實際應用	481
第21章	時間數列分析	495

6 目 錄

21-1	概 述	495
21-2	長期趨勢	497
21-3	季節變動	497
21-4	隨機變動(自我相關)	501
21-5	預 測	502
21-6	循 環	504
21-7	譜相分析	507
21-8	自我相關如何影響長期趨勢及季節變動之估計	513
第22章	聯立方程式	525
22-1	簡 介	525
22-2	經濟上的例子	526
22-3	工具變數(IV)	530
22-4	聯立方程式的推定	532
第23章	指 數	539
23-1	物價指數	539
23-2	其他指數	542
23-3	實物上之探討	545
23-4	實質所得的指數比較	547
第24章	抽樣設計	559
24-1	分層抽樣	559
24-2	其他抽樣設計	564
24-3	結 論	565
第25章	賽局理論	569
25-1	零和賽局	569
25-2	零和賽局的混合策略	572
25-3	結 論	575

附 錄.....	581
I. 常對數表.....	583
II. (a)隨機數字表.....	586
(b)常態隨機數表(標準化).....	587
III. (a)二項係數, $\binom{n}{x}$	588
(b)二項機率(逐項).....	589
(c)二項機率(累加).....	591
IV. 標準常態機率(累加).....	593
V. 學生 t 臨界點.....	594
VI. (a) X^2 臨界點.....	595
(b)修正之 X^2 (C^2) 臨界點.....	596
VII. F 臨界點.....	597
VIII. Wilcoxin-Mann-Whitney 二樣本檢定.....	599
IX. Durbin-Watson 臨界點.....	601
X. 表 V 至 VII 之插補.....	603

第 1 章

導 論

使用統計就如同醉漢扶持燈柱一般——其目的是為

支持，而非照明。——

—— Andrew Lang

“統計”一詞，本意為一國人口和重要經濟資料的搜集。但如今，它已發展成一種科學分析方法，而廣泛地應用在社會科學和自然科學之中。我們以下面的例子來說明統計學的目的和方法。

1-1 例 子

每屆選舉，一些跑地方政治新聞的記者總喜歡預測誰將當選，有時他們更預測每位候選人的得票率。很顯然地，對每一位選民詢問他將投給誰是不可能的，唯一的方法是，記者先生們從選民中抽取幾十個人為樣本來調查，而後根據此樣本，求得一個好的樣本比率，定值來推估整個母體的比率。此即統計推論 (statistical inference) 或統計歸納法 (statistical induction) 的一個典型例子。本例中，即由觀察到的樣本特性，來推論未知母體的特性。

每一位記者都得承認他的推估帶有不確定性。欲確知母體為何，則必待到投票日將所有的選票點計完後方可得知。但是，若抽樣為公正且適當的話，樣本比率將相當地接近母體比率。於是，我們可用觀

察到的樣本比率 P 來推定未知母體比率 π ，如下：

$$\pi = P \pm \text{小的誤差} \quad (1-1)$$

現在有一些重要的問題，即“此一小的誤差多小？”和“我們如何確定我們是對的？”由於此等問題是本書的核心所在，故我們敘述其精確的公式如下（此段取自第 8 章，屆時我們會再詳細地加以解釋）：

在隨機抽樣時，對

$$\pi = P \pm 1.96 \sqrt{\frac{P(1-P)}{n}} \quad (1-2)$$

我們有 95 % 的信賴度。在此 π 和 P 分別為母體比率和樣本比率， n 為樣本大小。

在例示此一公式之前，我們再重複序言中所說過的話：本書中每一數字性的例子便是一個習題，讀者應主動地去解它，而非被動地讀它。因此，每一例題我們皆以問題的形式列出，以待你的求解；若你做不出，則可讀答案。必須記住的是，統計學不是一門光說不練的學問，單是看，別想會騎腳踏車，統計學亦是如此。你必須跳下水，而後方能激起一些水花。

【例 1-1】如上述地方選舉之例。設有甲，乙二位候選人參加競選，某記者經抽樣調查 2,000 名選民，得其中 760 名將投給甲，而 1,240 名將投給乙。求投票給甲候選人之母體比率 π 的 95 % 信賴區間。

【解】樣本大小 n 為 2,000，而樣本比率為：

$$P = \frac{760}{2,000} = .38$$

將此代入 (1-2)：

$$\pi = .38 \pm 1.96 \sqrt{\frac{.38(1-.38)}{2,000}}$$

$$\pi \simeq .38 \pm .02 \quad (1-3)$$

此意即該地區全體選民投票給甲候選人的比率，介於36%和40%的信賴度為95%。（註一）

本書的主要目的，乃是建立如（1-3）的信賴區間；而另一個目的，則為與其有關的假設檢定。現用同一例來說明。設有某熱心的記者宣稱甲候選人將贏得此次選舉，以數學方式來說，此假設可寫成 $\pi > .50$ 。若以方程式（1-3）所得的結果來下判斷，則棄却此一假設。一般而言，區間推定和假設檢定之間有此類非常密切的關係。

對方程式（1-2）應注意下列數端：

1. 推定值並非確定值，而是只有95%的信賴度。由於可能不幸抽得一個不好的樣本，因而我們必須承認做錯決策的機率。例如，此次選舉中選民欲投甲候選人的比率，事實上不到一半。雖然不大可能發生，但我們仍然可能由樣本做出甲候選人將贏得選舉的結論。在此情形下，由信賴區間（1-2）便做出錯誤的結論。由於此等壞運氣雖然可能但不大可能發生（有5%的不可能），故我們有95%的信賴度。

2. 當樣本大小 n 增加時，（1-2）中的容受誤差減小。若例1-1中的樣本大小增加到10,000個選民，而後計算得甲候選人的比率為.38，則其95%信賴區間將更精確：

$$\pi \simeq .38 \pm .01$$

此點可由直覺來說明：大樣本比小樣本包含有更多的情報，因此它能得更精確的結論。

3. 設95%的信賴度不夠，而欲有99%信賴度的結論，若此時沒辦法再抽取樣本，則只有減小精確度來增加我們的信賴度。在第8章中我們將證明，若信賴度由96%增加到99%，則公式（1-2）的係數必須由1.96增加到2.58；而得99%的信賴區間為：

註一：方程式（1-3）中的二意即“近似地等於”。

$$\pi \simeq .38 \pm .03$$

此一區間比 (1-3) 95% 的信賴區間還要寬，但精確度較低。由於我們欲正確的機率更高（即更確定），故必須犧牲一些精確度。必須了解的是，任何情形下的統計敘述，必帶有某些不確定性。

1-2 演繹法和歸納法

圖 1-1 說明了演繹法和歸納法間差異的源由。(a)中的演繹法乃由一般情形論證到特殊狀況——亦即由母體到樣本。而(b)中的歸納法則反其道而行——由特殊狀況推論到一般狀況，即由樣本推論到母體。方程式 (1-1) 所表示者即為歸納法，我們由樣本比率來對母體比率加以推論。但是，歸納法僅在先行研究簡單的演繹問題之後，方為可能。以方程式 (1-1) 來說，此一歸納敘述（由樣本比率推定母體比率）乃是一事前的演繹（prior deduction）（樣本比率似是接近於母體比率）。

本書第 3 章到第 7 章討論演繹法，其所講的即為機率理論，該部份討論的主題是“從一已知母體中抽出樣本，此樣本的行為將如何？此樣本是否為目標樣本？”在解決此等演繹問題之後，以後各章便集



圖 1-1 演繹法和歸納法的比較(a)演繹法（機率）