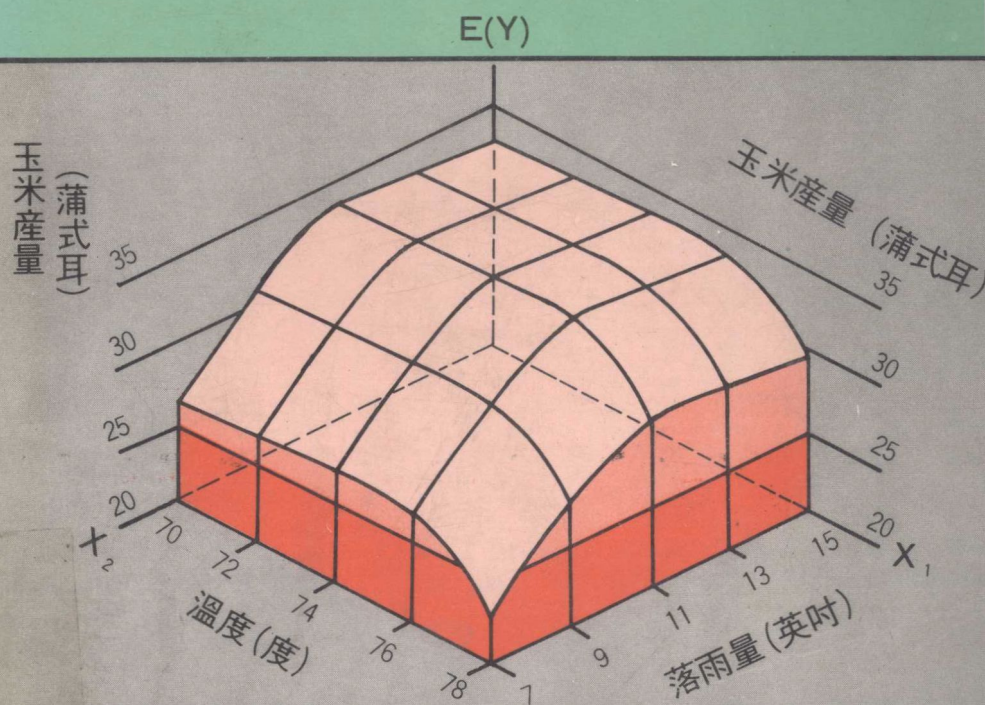


迴歸分析

—理論與應用—

(新增修訂本)

吳宗正 編著



$$E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

復文書局

迴歸分析

—理論與應用—

Applied Linear
Statistical Models

Regression Analysis of Variance
and Experimental Designs

John Neter

William Wasserman

吳宗正 編譯

(增修訂新版)

復文書局

江苏工业学院图书馆
藏书章

0212.1
W929

迴歸分析

—理論與應用—

版權所有



翻印必究

中華民國六十七年三月初版發行

中華民國六十九年九月增訂新版

中華民國七十一年九月增修訂三版

中華民國七十三年十月增修訂四版

全冊平裝280元 全冊精裝330元

著 作 者： 吳 宗 正

發 行 者： 吳 主 和

發 行 所： 復 文 書 局

地址：臺南市東門路421巷28號

門市：台南市林森路二段63號

電話：(06)2370003 · 2386937

郵政劃撥帳戶 0032104-6號

No.28. LANE421 DONG-MEN
ROAD TAINAN TAIWAN REPUBLIC
OF CHINA

TEL: (06)2370003 · 2386937

本書局經行政院新聞局核准登記發給
出版事業登記證局版台業字第0370號

一版序言

「線型統計模式」(Linear Statistical Models)如迴歸、變異數分析和實驗設計，在今日被廣泛的應用於管理經營、管理科學及社會科學中。要完善的用這些方法，需要徹底地了解以下諸理論學說及在實際生活領域中因應用這些模式而遭遇的實際問題。而本文首先所提出的「迴歸分析」(Regression Analysis)是一本基本應用的書，它有效的使理論與應用融合為一，避免過度的偏重於理論使之與應用產生分離，或只是偏重於應用的基本理論而忽略其理論之來源及根據，它同時亦是變異數分析及實驗設計的基本工具書。

本書的特色在於理論與應用並重，除了對模式的來源，證明做合理的詮釋，另外更舉實際例證配合模式的應用，而且在重要章節裡另闢“註解”與“評註”，對整個模式的理論及應用做更詳細和完整的分析及比較。近來年，由於工商業進步迅速，對於經營管理，市場營運，交通流量分析以及公司決策方面極待做全面有效的發展與配合，因此為達到經營合理的目標，須對未來的需求及供應，做事先的預測，以達到有效的控制。但是從事預測工作必須具有科學的方法及理論才能使結果預測精確，變異減小。因此有效地運用迴歸分析方法更能使預測工作得心應手，進行順利。

本書所介紹的迴歸分析實為預測方法的基礎，它是屬於計量方法的運用，所謂預測是對某事物未來可能演變的情形，事先予以推測。因此對於預測問題的處理，將依選擇的方法及技術設定適合的模式以配合資料，再從模式中計算相關的預測值，這類方法與技術在本書章節中成一系統合理的介紹及分析。

由於迴歸分析應用範圍廣泛，舉凡工商、經濟、管理領域皆能有效地使用，因此只要學過統計學科同學或工商界人士均能採用本書為參考書籍或自修之用。

本書編著工作承蒙成功大學統計系主任儲師全滋之多方鼓勵，以及諸位老師的多方指導，始能如期完成，在此表示衷心之謝意，又本

書的稿件承張華悌先生，丁承、詹文光、王秋月、張敏詩、蔡淑敏、黃娟貞、蔡慧蘭、黃女珍、蘇國樑、丁頌麒、洪金源及林貴雀諸位同學的抄錄及協助，統此致謝，而本書之出版承余振源同學的多方奔走及熱心協助，在此一併致謝。

編者學識淺短，且時間匆促，錯誤之處在所難免，尚祈學術界先進不吝賜教。

中華民國六十六年十二月

吳宗正 謹識

三 版 序 言

本書自六十六底開始發行，至今已經三版發售，感謝各界的鼓勵及支持，特藉再版之際，對愛護本書之人士，致衷心之謝意。

爲了使迴歸分析之應用更加廣泛，本書增加下列篇幅以供學者引證參考之用；(1)將重點部份打上 *號(2)對於書中需引證之部份增列參考資料以供對照，(3)爲配合電子計算機之應用，分別增加簡單迴歸，多重迴歸及多項式迴歸分析三部份的程式資料及系列分析過程，以供學者在學習及應用時對照參考之用。

本書三版發售承復文書局吳主和先生之支持及協助，在此特致謝意。

吳宗正 謹識於成大

中華民國七十一年九月一日

四版序言

本書自民國六十六年十二月初版發行以來，至今爲第四版發售，感謝各界的鼓勵及愛護，藉再次發行之際，特向支持愛顧本書之人士，致衷心地謝意。

迴歸分析在統計應用上非常地重要，爲使學者能廣泛地予以吸收及應用，本書重新將內容加以編排，以使學者易於瞭解，另外亦增列一些參考資料及相關的電腦程式設計以供學者參考引證之用，同時本書每章所附之習題內容極爲精闢，學者應將習題充份加以演習，以收理論及實務之配合。

本書四版發售承復文書局吳主和先生之鼎力支持及協助，在此特致謝意。

吳宗正 謹識於成功大學

中華民國七十三年十二月三十日

再版序言

自迴歸分析一書初版發行以來，承蒙各界之關注及支持，未及兩年即銷售一空。使編者引以爲榮，特藉再版之際，對愛護本書之人士，特致謝忱。

爲了使原書更完整，除了錯誤之處核校外，還另在書後附加參考資料一篇，以供閱讀本書時之參考引証，雖曾仔細費心去校正，難免還有不能顧及之處，尙祈專家先進賜予指正毋任感荷。

本書之再版承復文書局吳主和先生之支持及余振源同學之協助在此特致謝意。

又在校稿期間承摯友賢彬小姐之鼓勵及協助，謹致深摯地謝意！

吳宗正 謹識

中華民國六十九年三月十日

目 錄

第一章	機率與統計的基本原理	1
1 - 1	求和及乘法運算	1
1 - 2	機率	1
1 - 3	隨機變數	16
1 - 4	常態機率分配和相關之分配	27
1 - 5	統計估計	32
1 - 6	常態母羣體平均數之推定	37
1 - 7	兩常態母羣體平均數之比較	40
1 - 8	常態母羣體變異數之推定	43
1 - 9	兩常態母羣體變異數之比較	45

第一部份：基本迴歸分析

第二章	一個自變數之線型迴歸	49
2 - 1	兩變數間之關係	49
2 - 2	迴歸模式及其運用	53
2 - 3	誤差項分配之迴歸模式	63
2 - 4	迴歸函數之估計	67
2 - 5	誤差項變異數 σ^2 之估計	79
2 - 6	常態誤差迴歸模式	84
2 - 7	計算機輸入與輸出	88
第三章	迴歸分析之推定	93
3 - 1	關於 β_1 之推定	93
3 - 2	β_0 之推定	104
3 - 3	建立 β_0 和 β_1 推定必須考慮之因素	107
3 - 4	$E(Y_h)$ 之區間估計	109
3 - 5	新觀察值之預測	115
3 - 6	迴歸分析在應用上應考慮之因素	123
3 - 7	X 為隨機之情況	124

3 - 8	迴歸分析中之變異數分析法	125
3 - 9	迴歸模式中X與Y變數關係程度之測度	142
3 - 10	計算機輸出	146
第四章	模式的傾向性與矯正方法	152
4 - 1	剩餘值	152
4 - 2	剩餘值之圖形分析	154
4 - 3	剩餘值之檢定	167
4 - 4	綫型的F檢定	168
4 - 5	矯正方法	179
4 - 6	經由轉換之綫型迴歸函數	181
4 - 7	變異數穩定之轉換	189
第五章	迴歸分析研討—I	200
5 - 1	β_0 與 β_1 之聯合估計	200
5 - 2	迴歸綫之信賴界限帶	210
5 - 3	平均數 $E(Y_h)$ 之群組估計	214
5 - 4	新觀察值之群組預測區間	217
5 - 5	通過原點之迴歸綫	219
5 - 6	兩種迴歸綫之比較	224
5 - 7	測度誤差之影響	232
5 - 8	X水準之抉擇	236

第二部份：一般迴歸和相關分析

第六章	簡單迴歸分析之矩陣方法	244
6 - 1	矩陣	244
6 - 2	矩陣加法與減法	250
6 - 3	矩陣乘法	251
6 - 4	矩陣之特殊形式	256
6 - 5	綫型相依和矩陣之階次	258
6 - 6	反矩陣	259
6 - 7	矩陣之一些基本定理	265

6-8	隨機向量和矩陣	266
6-9	簡單綫型迴歸模式之矩陣表示法	270
6-10	迴歸參數之最小平方估計法	273
6-11	變異數因子之分析	276
6-12	迴歸推定方法	278
6-13	加權最小平方方法的矩陣表示法	285
6-14	剩餘值(殘差)的矩陣表示法	286
第七章	多重迴歸	294
7-1	多重迴歸模式	294
7-2	用矩陣項表示一般綫型迴歸	306
7-3	最小平方估計值	308
7-4	變異數因子之分析	308
7-5	迴歸參數之推定	312
7-6	平均數對應值之推定	314
7-7	新觀察值之預測	317
7-8	實例一兩個自變數之多重迴歸分析	319
7-9	多重迴歸模式分析之有關問題	336
7-10	多重迴歸模式之檢定假設	348
7-11	偏判定係數	356
7-12	一般線性檢定法的矩陣符號格式	359
7-13	標準化之迴歸參數	362
第八章	多項式迴歸	369
8-1	多項式迴歸模式	369
8-2	例一：一個自變數	375
8-3	例二：二個自變數的多項式迴歸模式	386
8-4	估計二次迴歸函數之極大值或極小值	391
8-5	多項式迴歸之詳論	393
第九章	指示變數	399
9-1	一個獨立屬質變數的模式	399
9-2	包含交互影響之模式	407

9 — 3	複合模式之論說.....	413
9 — 4	獨立指示變數之其他用途.....	419
9 — 5	使用獨立指示變數應考慮之因素.....	425
9 — 6	相依指示變數.....	428
9 — 7	相依指示變數之線型迴歸.....	432
9 — 8	羅吉斯對應函數.....	438

第十章 迴歸分析研討—II 451

10 — 1	轉換參數法——改善計算精確度的方法	451
10 — 2	多重共線性之本質及問題	458
10 — 3	變異數膨脹因素及其他偵測多重共線性的方法	470
10 — 4	山脊型迴歸法及其他對多重共線性發生時的補救方法	475
10 — 5	極端觀察值的辨認	485
10 — 6	強勢觀察值的辨認及補救方法	495

第十一章 最佳自變數的搜尋 506

11 — 1	問題的本質	506
11 — 2	例證	508
11 — 3	所有可能迴歸搜尋程序.....	510
11 — 4	分步迴歸法	522
11 — 5	搜尋程序的執行	531

第十二章 時間數列資料的自我相關性..... 540

12 — 1	序言.....	540
12 — 2	自我相關性的問題.....	540
12 — 3	第一階次自我迴歸模式.....	544
12 — 4	自我相關性的德賓——華德森 (Durbin-Watson) 檢定	547
12 — 5	迴歸參數的推估.....	552

第十三章 非線性迴歸模式 565

13 — 1	非線性迴歸模式之意義及應用——例證(1).....	565
13 — 2	非線性迴歸參數之最小平方推估法.....	571

13—3	非綫性迴歸參數之推定	586
------	------------	-----

13—4	學習曲綫之例證(2)	591
------	------------	-----

第十四章	常態相關模式	601
-------------	---------------	------------

14—1	迴歸和相關模式間的區別	601
------	-------------	-----

14—2	二元常態分配	602
------	--------	-----

14—3	條件推定	606
------	------	-----

14—4	ρ_{12} 的推定	611
------	-----------------	-----

14—5	多元變量常態分配	616
------	----------	-----

參考資料	627
------	-----

附 表	677
-----	-----

參考書目	693
------	-----

第一章 機率及統計之基本定理

1-1 求和及乘法運算

求和運算：

$$(1) \sum_{i=1}^n Y_i = Y_1 + Y_2 + \cdots + Y_n$$

求和運算中之一些重要性質：

$$(1) \sum_{i=1}^n K = n K \quad K \text{ 是常數}$$

$$(2) \sum_{i=1}^n (Y_i + Z_i) = \sum_{i=1}^n Y_i + \sum_{i=1}^n Z_i$$

$$(3) \sum_{i=1}^n (a + c Y_i) = n a + c \sum_{i=1}^n Y_i, \quad a \text{ 和 } c \text{ 是常數}$$

二重求和運算符號 $\sum \sum$ 其定義如下：

$$(1) \sum_{i=1}^n \sum_{j=1}^m Y_{ij} = \sum_{i=1}^n (Y_{i1} + Y_{i2} + \cdots + Y_{im}) \\ = Y_{11} + Y_{12} + \cdots + Y_{1m} + \cdots + Y_{2m} + \cdots + Y_{nm}$$

二重求和運算之重要性質如下：

$$(2) \sum_{i=1}^n \sum_{j=1}^m Y_{ij} = \sum_{j=1}^m \sum_{i=1}^n Y_{ij}$$

乘法運算：

乘法運算式子 π ，其定義如下：

$$(3) \prod_{i=1}^n Y_i = Y_1 \cdot Y_2 \cdots Y_n$$

1-2 機率

* 機率法則及概念

在從事實驗研究時由於實驗對象及環境之變因，須涉及機率之應用，因此就下列諸項問題加以研討：

- (1) 實驗研究包含實驗之單位，例如男性、豬仔、土地、工廠、可銷售貨品等，它們是依某些計劃（包含偶發因素）在研究主旨下抽樣或製造而得，對某一計劃它們不但由隨機抽樣而得，而且它們的存在、情況、態度等在實驗過程中須依賴機率因素。
- (2) 樣本實驗單位之測度值是隨機變數，因為(1)及測度誤差之原因。
- (3) 對於實驗資料之任何明顯和有效的詮釋須多限制於某些不確定之變異，而這些不確定變異之程度必須設法加以測度及表達。

上述三項皆涉及機率因素，而機率即為測度不確定變異和明確有效地表達變異的工具，幸運地，數理機率之理論自從 1900 年來已經發展地很完整，統計理論即包含數理機率、隨機化、估計及統計假設檢定之觀念，而這些觀念亦為吾人在做實驗研究時所關心之主題。

機率所牽涉之觀念衆多，但是吾人將集中於自然的相對次數機率觀念上，它可以從實驗中模擬而得，為方便計，茲用銅幣和骰子做為引介機率基本觀念之工具，畢竟它們在實驗研究中佔著重要之地位。

若投擲一枚公平的銅幣，讓其落於平面上，無論結果是“正”或“反”在上面，大部份之觀察者認為實驗之結果（outcome）“正”“反”出現之可能性應該相等，但是某些觀察者會認為不盡然相等，例如他認為擲一枚銅幣 10 次，“正”“反”各出現 5 次之機會即不會相等，可是，實驗次數增加至 10000 次，任何觀察者皆會為其正反各出現 5000 次之結果而震驚，因此，可以顯示正反出現之機率各為 $\frac{1}{2}$ ，上述之實驗即說明當實驗次數增加時，正反兩面出現

之機會愈接近，投擲次數愈少，則出現正反面之次數與總次數之比例偏差 (bias) 愈大，例如在 5000 次投擲中出現正面之機率為 $0.33 = P_H$ ，其結論可以說明此銅幣有偏差，此時投擲反面之機率為 $P_T = 1 - P_H = 0.67$ 。

上述之實驗說明一件事，即對可描述的事件經過實驗之過程，吾人可以測度其不同結果之相對次數，使其機率之和為 1，因此，實驗所有可能之結果所形成之集合，稱之為樣本空間，其和為 1，樣本空間 (sample space) 有下列諸項性質：

- (1) 一實驗之所有可能結果若在空間可以以樣本點 (sample) 表示，其集合稱為樣本空間。
- (2) 每一樣本點有一發生之機率，其測度機率是依出現之次數與總次數之相對次數比例表示。
- (3) 樣本點之機率和等於 1 (100 %)，若為不偏銅幣之實驗，投一次其出現正反面之相對次數各為 $1/2$ ，其總和等於 1，若實驗一偏態的銅幣，在空間中仍然可得正、反兩點之相對次數，例如 $P_H = 0.33$ ， $P_T = 0.67$ ，其總和亦等於 $P_T + P_H = 1$ 。

茲從事一實驗，投擲 4 枚銅幣 1 次，總共有 $2^4 = 16$ 次可能之結果，設 4 枚銅幣分別為 5 元 (p)、1 元 (n)、1 角 (d)、1 分 (q)，其實驗結果如下：

p n d q	出現正面之次數	事件次數	相對次數—機率
HHHH	4	1	$1/16$
HHHT	3	4	$4/16$
HH'TH			
H'THH			
HTHH			

p n d q	出現正面之次數	事件次數	相對次數—機率
T H H H	2	6	$6/16$
H H T T			
H T H T			
T H H T			
H T T H			
T H T H			
T T H H	1	4	$4/16$
H T T T			
T H T T			
T T H T			
T T T H	0	1	$1/16$
T T T T			

若投擲銅幣為一公平之實驗，則投擲 4 枚銅幣之所有可能結果為 16 次，而在任一可能結果出現之機率應為 $\frac{1}{16}$ 。

上述每一實驗結果表示一樣本點，依其出現之次數與總次數相比，則得到相對次數之機率，右項一欄其各機率之總和應為 1。下列將引介一些機率常用之法則以供參考：

* (一) 加法定理——和事件之探討

- (1) 兩事件中有一事件發生或同時發生之機率定理稱為加法定理。對一樣本空間之任意兩事件 A_1 及 A_2 其加法定理敘述如下：

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

其中， $P(A_1 \cap A_2) \neq 0$