

醫學類別資料分析

Categorical Data Analysis in Medicine

林建甫

Jeff Chien-Fu Lin, MD. PhD.

醫學類別資料分析 = Categorical data analysis in medicine

/ 林建甫著. - 一版. - 台北市 : 雙葉書廊, 2020. 08

594 面; 19*26 公分

ISBN 978-957-9096-90-4 (平裝)

1. 生物統計 2. 醫學 3. 統計分析

306.13

109012489

©2020 林建甫

書名: 醫學類別資料分析

著者: 林建甫

著作權與版權所有人: 林建甫

出版社: 雙葉書廊有限公司

發行人: 張福隆

地址: 台北市羅斯福路三段 269 巷 12 號 1 樓

<http://www.yehyeh.com.tw>

電話: 886-2-2368-4198

傳真: 886-2-2365-7990

出版日期: 西元 2020 年 8 月 一版一刷

ISBN: 978-957-9096-90-4 (平裝)

本書著作權與版權屬於著作者所有, 受中華民國著作權法保護, 未經著作者書面同意, 不得以任何形式轉載, 複印, 翻印或重製。

作者 林建甫

Jeff Chien-Fu Lin, MD. PhD.

cflin@mail.ntpu.edu.tw

<https://jefflinmd.com>

經歷 巨匯數據科技股份有限公司 統計顧問

美國范德堡大學醫學中心 計量科學中心 生物資訊與癌症研究 訪問學者

智策市場研究顧問公司 統計顧問

台北大學 統計系 助理教授

台北榮民總醫院 生物統計顧問

台北市 中山醫院 骨科醫師

台北醫學大學 萬芳醫院 骨科醫師

高雄長庚醫院 骨科醫師

台北榮民總醫院 骨科醫師

學歷 美國密西根大學 生物統計研究所 碩士, 博士

高雄醫學大學 醫學系 醫學士

序

在醫學研究中，類別資料是一種常見的資料類型，醫學研究結果的測量通常是類別變數，例如，臨床試驗在受試者服用抗骨質疏藥物，追蹤數年後研究是否可降低骨折發生。又例如，第二期癌症臨床試驗在試驗藥物治療後的療效分成：完全反應，部分反應，穩定，持續進展等，追蹤後研究試驗藥物是否有顯著療效。這些研究與分析類別變數的統計方法為類別資料分析，這是一個籠統定義的統計名詞。類別資料也經常出現在不同的研究領域中，例如，社會學研究抽樣調查，市場研究，經濟計量的離散選擇模型等。類別資料分析相關書籍中，有許多可做為參考，經典書籍如 Agresti (2012, 2018), McCullagh 與 Nelder (1989), Collett (2003), Cox 與 Snell (1989) 等等，個別統計軟體也出版其專用書籍，例如，SAS 的 Stokes, Davis 與 Koch (2012), STATA 的 Hardin 與 Hilbe (2018), R 的 Friendly 與 Meyer (2016) 等等。這些書籍同時納入醫學，社會學研究或其他領域研究的例題，因此對長期從事臨床醫學研究的醫師，較難以快速閱讀與理解，因此本書探討的類別資料分析方法主要以分析臨床醫學研究為主軸，避免納入過多的其他研究領域的議題。

本書主要寫作對象為從事資料分析的統計人員與從事臨床醫學研究的醫師，本書所使用的數學內容僅包含加法，乘法，與簡單之微積分，這些內容等同於在台灣的高中與大一的基礎數學程度。從事臨床醫學研究的醫師閱讀本書，最好具備基本醫學統計的知識，並建議臨床醫學研究的醫師閱讀本書時，以閱讀文字為主，方程式演算過程為輔，主要須了解書中文字與演算方程式之結果，與方程式背後在資料分析實務上之意義，對每一種分析方法的使用時機與其優缺點，這樣則會有助於判讀醫學文獻與選擇適當的分析方式分析手中的研究資料。

本書共 17 章，第 1 – 2 章介紹類別資料基礎統計。第 3 – 7 章討論列聯表分析，包含二維列聯表，分層列聯表，配對列聯表與醫學診斷工具，以及多維列聯表。第 8 章介紹類別資料抽樣分配與第 9 章介紹廣義線性模型理論。第 10 – 13 章討論邏輯斯迴歸模型，邏輯斯迴歸診斷，條件邏輯斯迴歸與多項式邏輯斯迴歸。第 14 章討論事件發生

速率之布瓦松迴歸模型, 第 15 章介紹重覆測量與相關性資料之基本分析, 第 16 – 17 章討論資料內的所有變數皆為類別變數之多維列聯表與對數線性模型等相關議題. 本書以應用在醫學類別資料分析為主, 對於單變量分析方法中, 若現今常見之統計軟體可以執行之分析方法, 則儘可能提出討論. 本書中所有的例題分析與圖表建構, 皆是以使用 R 或是 SAS 統計軟體執行運算, 使用這二種統計軟體主要是因為作者的偏好與熟悉度, 本書中多數的例題分析與圖表建構, 也可以使用其它統計軟體, 如 STATA, SPSS 等執行運算. 本書如同其它書籍, 不可能討論所有類別資料分析的議題, 在本書中割捨許多不常出現在臨床醫學資料分析的內容, 本書也割捨多變量類別資料與縱向類別資料分析中的複雜理論與模型. 最後, 因本人才疏學淺, 以致本書有所錯誤或疏漏, 尚請諸位先進與賢達不吝指教.

本書是作者在 2002–2020 年間寫作, 本書的起源約在 2002–2004 年, 作者在台北大學統計系擔任兼任教師, 以電子書型式作為學士班與碩士班講授類別資料分析課程的講義. 2002 年原稿使用英文寫作, 原始內容包括不同領域的類別資料分析, 隔年並將大部份內容改寫與修訂成中文. 在 2006 年後因改授其它課程而中斷寫作, 在 2007–2018 年間, 斷斷續續講授醫學統計相關課程, 同時也斷斷續續改寫部分內容, 在 2008, 2015, 2018, 2019 年都有較大幅度修改, 由於未繼續在統計系講授類別資料分析, 因此在 2008 年以後修改本書時, 都是著重在臨床醫學研究資料分析的相關議題, 而排除醫學研究以外的應用, 最後於 2020 年暑期完成寫作計畫.

作者長期從事醫學統計諮詢工作與期刊論文的統計審核, 嘗試結合作者個人臨床醫學研究與醫學統計諮詢經驗, 以醫學與統計雙方都較能了解的語言與文字, 將醫學類別資料分析之統計方法整理成書. 希望藉由本書之寫作, 在統計人員與臨床醫師之間, 提供一個溝通的橋樑, 並共同分享類別資料分析的經驗.

人生充滿著意外與驚喜, 在偶然的機緣下, 進入到醫學與統計的世界, 先後受到醫學與統計兩個學門的師長, 悉心指導, 感激在心. 本書的完成, 要感謝美國密西根大學生物統計系給我機會, 進入到統計的世界. 感謝讓我有機會進入生物統計系就學的教授 Michael Boehnke 博士, 感謝我的指導教授 Robert A. Wolfe 博士; 感謝在求學期

間提供我助學金的教授 Richard G. Cornell 博士, Judith Bromberg 博士, M. Anthony Shorck 博士. 這些年來我有機會繼續從事跨領域的研究與工作, 在統計工作上感謝國立台北大學統計系的前輩與同事協助, 感謝中國醫藥大學梁文敏博士給予作者在研究上長期的支援與合作. 在臨床工作上受到台北市萬芳醫院骨科何為斌主任, 郭宜潔主任, 以及骨科所有同仁的協助, 特此致謝, 也特別感謝台北中山醫院吳濬哲院長與台北市仁愛醫院骨科呂才學主任給予作者長期的支持與鼓勵, 本書的排版是利用吳聰敏教授, 吳聰慧先生, 與翁鴻翎先生開發出來的 `cwTeX` 系統排版軟體所完成, 謹此致謝. 本書能夠付梓, 感謝雙葉書局編輯部門的協助. 最後, 感謝在這段期間, 鳳凰, 湘宜, 靖庭的犧牲與支持. 讓我有機會繼續從事跨領域的工作.

林建甫 謹誌于
台北大學統計學系
2020 年 8 月

目錄

第 1 章 類別資料分析導論	1
1.1 研究資料與變數	1
1.2 醫學研究設計分類	7
1.2.1 橫斷性研究	8
1.2.2 回溯性研究	9
1.2.3 前瞻性研究	10
1.2.4 世代研究	11
1.2.5 臨床試驗	14
1.3 離散型機率分配	17
1.3.1 離散均一分配	17
1.3.2 伯努利分配	18
1.3.3 二項分配	20
1.3.4 幾何分配	22
1.3.5 超幾何分配	24
1.3.6 布瓦松分配	24
1.3.7 多項式分配	28
1.3.8 負二項分配	29
第 2 章 事件發生分率與速率之推論	33
2.1 事件發生機率	34
2.2 事件發生速率	36
2.3 差異與比率	38

2.4	事件發生分率推論: 單一樣本與二項分配	40
2.4.1	單一樣本事件發生分率: 二項分配近似檢定	41
2.4.2	單一樣本事件發生分率: 二項分配信賴區間	42
2.4.3	單一樣本事件發生分率: 二項分配精確檢定	44
2.4.4	精確檢定與中間 p -值	45
2.5	二組樣本事件發生分率: 二項分配近似 Z 檢定	48
2.6	事件發生速率與布瓦松分配	56
2.7	一組樣本事件發生速率: 布瓦松分配與推論	60
2.8	二組樣本事件發生速率: 布瓦松分配與推論	65
2.9	事件發生速率: 標準化	67
2.9.1	直接標準化	70
2.9.2	間接標準化	73
2.10	資料分析練習	76
第 3 章 列聯表分析		79
3.1	一維列聯表分析: 卡方檢定配適度	80
3.2	二維列聯表簡介	83
3.3	前瞻性研究與二維列聯表	86
3.4	橫斷性研究與二維列聯表	99
3.5	回溯性研究與二維列聯表	102
3.6	可歸因危險	109
3.7	2×2 列聯表: 關聯性檢定	112
3.7.1	皮爾森卡方檢定	114
3.7.2	概似比檢定	115
3.7.3	費雪精確檢定	116
3.8	$R \times C$ 列聯表	124
3.9	Cochran-Armitage 趨勢檢定: $2 \times K$ 列聯表	128
3.10	資料分析練習	132

第 4 章	分層列聯表分析	143
4.1	分層列聯表	143
4.2	整體關聯性: Mantel-Haenszel 檢定	148
4.3	齊一性卡方檢定	149
4.4	共同勝算比估計	151
4.5	辛普森詭論	154
4.6	資料分析練習	157
第 5 章	二維配對列聯表	161
5.1	2×2 二維配對列聯表	162
5.2	配對比較二組差異: McNemar 檢定	165
5.3	可靠度與一致性	169
5.4	資料分析練習	176
第 6 章	醫學診斷工具分析	181
6.1	敏感度與特異度	182
6.2	ROC 曲線分析	190
6.3	資料分析練習	199
第 7 章	多維列聯表	203
7.1	三維列聯表	204
7.2	完全獨立模型	206
7.3	聯合獨立模型	208
7.4	條件獨立模型	209
7.5	同質關聯模型	211
7.6	建模對策	213

第 8 章 類別資料抽樣分配	215
8.1 布瓦松分配與布瓦松抽樣	216
8.2 多項式分配與式多項式抽樣	217
8.3 多項式分配與獨立多項式抽樣	218
8.4 多項式分配與抽樣分配	220
8.5 類別資料配適統計模型	225
8.6 配適飽和模型與參數估計	226
8.7 概似函數, 分數向量以及資訊矩陣	227
8.8 統計模型參數估計式之大樣本漸近理論	229
8.9 多項式分配方格機率估計式之大樣本漸近理論	231
8.10 統計模型參數估計式與方格機率估計式之漸近分配	233
8.11 皮爾森殘差的漸近分配	235
8.12 皮爾森卡方分配的漸近分配	236
8.13 概似比檢定統計量的漸近分配	237
8.14 離差殘差	239
第 9 章 廣義線性模型理論	241
9.1 廣義線性模型	242
9.2 廣義線性模型的成分	243
9.3 指數族分配	245
9.4 連結函數	247
9.5 充分統計量與典型連結	249
9.6 廣義線性模型: 概似方程式	250
9.7 廣義性模型: 參數估計	252
9.8 典型連結的簡化估計	258

9.9	參數推論: 假說檢定	259
9.10	參數推論: 信賴區間	261
9.11	配適度檢定	263
9.12	檢定巢狀模型	265
9.13	模型配適: 過程與原則	267
9.14	廣義線性模型殘差	268
第 10 章	邏輯斯迴歸模型	271
10.1	線性機率模型	272
10.2	邏輯斯迴歸模型簡介	275
10.3	簡單邏輯斯迴歸模型	276
10.4	複邏輯斯迴歸模型	285
10.5	參數估計與推論	291
10.6	配適度檢定	295
10.7	檢定巢狀模型	297
10.8	檢定二個模型	298
10.9	事件發生機率之推論與預測	304
10.10	模型選擇	308
10.10.1	模型選擇標準與模型選擇程序	309
10.10.2	預測誤差與分類評估	311
10.11	資料分析練習	318
第 11 章	邏輯斯迴歸: 模型診斷	325
11.1	殘差	326
11.2	影響值	331
11.3	殘差分析	333
11.4	共變數的最佳函數形式: 加入自變數圖	338

11.5	共變數的最佳函數形式: 局部殘差圖	340
11.6	共變數的最佳函數形式: 變數轉換	341
11.7	檢視變異數函數	342
11.8	調整適度配缺失	343
11.9	變異數過度離散	344
11.10	貝塔-二項式模型	355
11.11	資料分析練習	360
第 12 章	條件邏輯斯迴歸	363
12.1	配對設計	364
12.2	條件邏輯斯迴歸: 1:1 配對	364
12.3	條件邏輯斯迴歸: 1:M 配對	368
12.4	條件邏輯斯迴歸模型: M:N 配對	373
12.5	資料分析練習	375
第 13 章	多項式邏輯斯模型	377
13.1	基線類別邏輯斯模型	378
13.2	離散選擇模型	384
13.3	順序反應變數與與累積連結模型	387
13.4	順序反應變數與累積邏輯斯迴歸模型	390
13.5	相鄰類別邏輯斯模型	394
13.6	順序反應變數與連續比邏輯斯模型	396
13.7	資料分析練習	399
第 14 章	布瓦松迴歸模型	403
14.1	對數線性迴歸模型簡介	404

14.2	事件發生速率與布瓦松迴歸模型	405
14.3	簡單布瓦松迴歸模型	406
14.4	多變數布瓦松迴歸模型	410
14.5	參數估計與推論	413
14.6	配適度檢定	423
14.7	變異數過度離散	429
14.8	負二項分配迴歸模型	431
14.9	離散過低	435
14.10	零的方格計數	435
14.11	資料分析練習	443
第 15 章	重覆測量與相關性資料	449
15.1	相關性資料	450
15.2	相關性類別資料分析模型	453
15.3	邊際模型	454
15.4	廣義線性混和模型	464
15.5	資料分析練習	470
第 16 章	列聯表與對數線性模型	473
16.1	一維列聯表與對數線性模型	473
16.2	二維列聯表與對數線性模型	475
16.2.1	二維列聯表: 獨立模型	477
16.2.2	二維列聯表: 飽和模型	478
16.2.3	參數解釋	479
16.3	三維列聯表與對數線性模型	481
16.3.1	三維列聯表對數線性模型	482
16.3.2	參數估計	485
16.3.3	參數解釋	487

16.4	階層模型	489
16.5	邏輯斯模型與對數線性模型	490
16.6	資料分析練習	496
第 17 章	配對列聯表與廣義線性模型	497
17.1	對稱模型	498
17.2	準對稱模型	499
17.3	準獨立模型	500
17.4	檢定邊際同質性	501
17.5	順序條件對稱模型	502
17.6	順序準關聯模型	503
17.7	檢定順序變數的邊際同質性	505
17.8	測量的一致性與測量相關性對數線性模型	506
17.9	配對比較與 Bradley-Terry 模型	509

表目錄

1.1	糖尿病與人工全膝關節置換手術研究: 部分個體資料	4
1.2	糖尿病與人工全膝關節置換手術研究: 變數說明	5
1.3	肺癌與二手菸	5
1.4	肺癌與酗酒	6
2.1	2 組母體 與 2 組獨立樣本: 分率檢定符號	49
2.2	二個獨立樣本假說: 比較母體的事件發生機率	50
2.3	飲食攝取糖量與糖尿病	64
2.4	某地區粗死亡率與特定年齡死亡率	68
2.5	Sweden 與 Panama 某一年的死亡率	69
2.6	標準化死亡率: 參照人口結構	70
2.7	Sweden 死亡率: 直接標準化	72
2.8	間接標準化死亡率: Panama 與 Sweden 死亡率	74
2.9	偏頭痛臨床試驗: 變數說明	77
2.10	NHS 世代研究: 口避孕藥與乳癌	78
3.1	孟德爾豌豆遺傳實驗: 觀察資料與期望值	80
3.2	DM-TKR 實驗組別與感染: 一維列聯表人數摘要	84
3.3	DM-TKR 實驗組別與感染: 2×2 列聯表摘要	84

3.4	2 × 2 的列聯表: 觀測個數摘要	85
3.5	2 × 2 的列聯表: 期望個數摘要	85
3.6	前瞻性研究初期: 2 × 2 的列聯表摘要	87
3.7	前瞻性研究結束: 2 × 2 的列聯表摘要	87
3.8	PHS 臨床試驗: 阿斯匹靈與心肌梗塞	88
3.9	阿斯匹靈與心肌梗塞: 條件機率估計	91
3.10	橫斷性研究: 子宮內避孕器 (IUD) 使用時間與不孕症研究	100
3.11	橫斷性研究: 2 × 2 的列聯表摘要	100
3.12	回溯性研究初期: 2 × 2 的列聯表摘要	103
3.13	回溯性研究結束: 2 × 2 的列聯表摘要	103
3.14	吸菸與肺癌	104
3.15	回溯性研究: 2 × 2 的列聯表機率摘要	105
3.16	回溯性研究: 2 × 2 的列聯表整體摘要	106
3.17	疫苗的臨床試驗	111
3.18	轉移性乳癌與高血鈣臨床試驗	112
3.19	2 × 2 列聯表: 觀測個數摘要	113
3.20	2 × 2 的列聯表: 期望個數摘要	113
3.21	費雪精確檢定: 假設在 2 × 2 列聯表中 $(1,1) = 0$	117
3.22	費雪精確檢定: 假設在 2 × 2 列聯表中 $(1,1) = 1$	117
3.23	費雪精確檢定: 假設在 2 × 2 列聯表中 $(1,1) = a$	118
3.24	費雪精確檢定: 假設在 2 × 2 列聯表中 $(1,1) = k - 1$	118
3.25	費雪精確檢定: 假設在 2 × 2 列聯表中 $(1,1) = k$	118
3.26	橫斷性研究: 宗教信仰與教育	119
3.27	回溯性研究: 半髖人工關節置換手術後髖關節脫臼與感染	121
3.28	鼻咽癌回溯性臨床研究	123
3.29	R × C 列聯表	125

3.30	$R \times C$ 表: 關節炎與止痛藥物治療的臨床試驗	125
3.31	$2 \times K$ 列聯表: 大學體育系就讀年級與運動傷害	128
3.32	$2 \times K$ 列聯表趨勢檢定: 變數 Y 中 K 個類別水準與分數	129
3.33	橫斷性研究: 性別與癌症	132
3.34	病例對照研究: 維生素 K 與癌症	133
3.35	回溯性研究: 染幽門螺旋桿菌與胃癌	134
3.36	Hismanal 臨床試驗	135
3.37	回溯性病例對照研究: 沙利竇邁與嬰兒四肢缺陷	136
3.38	抗凝血劑與血管支架臨床試驗	136
3.39	乾眼症臨床試驗	137
3.40	AIDS 疫苗試驗: mITT 分析	137
3.41	AIDS 疫苗試驗: ITT 分析	138
3.42	AIDS 疫苗試驗: PP 分析	138
3.43	胰臟癌與年紀	139
3.44	梨狀肌症候群臨床試驗: 變數說明	141
4.1	分層回溯性研究: 性別, 心電圖與冠狀動脈疾病	144
4.2	性別與冠狀動脈疾病: 邊際列聯表	145
4.3	心電圖與冠狀動脈疾病: 邊際列聯表	145
4.4	心電圖與冠狀動脈疾病: 男性部分列聯表	146
4.5	心電圖與冠狀動脈疾病: 女性部分列聯表	146
4.6	第 i 層: 事件 (疾病) 與危險因子之 2×2 列聯表, $i = 1, 2, \dots, K$	147
4.7	預防性抗生素與泌尿系統感染: 邊際列聯表	155
4.8	預防性抗生素與泌尿系統感染: 分層條件列聯表	155
4.9	高鈣飲用水與骨折	157
4.10	肺炎: 多中心臨床試驗	158

4.11	WHS 臨床試驗	159
5.1	骨關節炎疼痛控制臨床試驗: 2×2 獨立列聯表	162
5.2	骨關節炎疼痛控制臨床試驗: 2×2 配對列聯表	164
5.3	McNemar 檢定: 2×2 相依列聯表	165
5.4	心血管狹窄數目: 影像判讀的一致性	170
5.5	配對資料: 一致性分析	171
5.6	κ 統計量: 一致性等級	173
5.7	營養調查資料: 配對 2×2 列聯表	174
5.8	配對 2×2 列聯表: kappa 統計量差異	175
5.9	子宮內膜癌與雌激素配對病例對照研究	176
5.10	胰島素幫浦治療不穩定血糖: 治療前後比較	177
5.11	十二指腸潰瘍臨床試驗: 2×2 配對列聯表	177
5.12	眼睛過敏: 變數說明	178
5.13	頸椎畸形: 2 位臨床醫生診斷	178
5.14	冠狀動脈造影診斷可靠度研究	179
6.1	醫學檢驗結果與疾病	183
6.2	醫學檢驗與假說檢定	184
6.3	敏感度 (sensitivity) 與 特異度 (specificity): 2×2 方正列聯表	185
6.4	盛行率, 敏感度與特異度: 2×2 方正列聯表	185
6.5	乳癌篩檢與血中雌二醇濃度: 病例對照研究	188
6.6	敏感度, 特異度 與 盛行率: 近似計算	190
6.7	乳癌篩檢與血中雌二醇濃度資料	191
6.8	乳癌篩檢: 以血中雌二醇濃度 30 pg/ml 為切點	196
6.9	不同血中雌二醇濃度為切點: 敏感度與特異度	196
6.10	不同血中雌二醇濃度為切點: PPV 與 NPV	197

6.11	血管攝影診斷工具資料: 變數說明	199
6.12	骨骼掃描與前列腺癌檢驗資料: 變數說明	200
6.13	細胞彈性蛋白 U+9176 檢驗資料: 變數說明	201
9.1	一般單變數的指數族分配與常用的連結函數	249
9.2	常見機率分配與典型連結	249
10.1	血壓與心血管疾病	274
10.2	血壓與心血管疾病: 簡單線性迴歸模型	275
10.3	骨質疏鬆臨床試驗	283
10.4	骨質疏鬆臨床試驗: 變數名稱與定義	283
10.5	簡單邏輯斯迴歸: 概似比檢定	284
10.6	簡單邏輯斯迴歸: 參數估計	284
10.7	醫材臨床試驗: 變數名稱與定義	288
10.8	醫材臨床試驗: 資料	288
10.9	邏輯斯迴歸: 參數估計	290
10.10	低新生兒體重研究: 變數說明	300
10.11	低新生兒體重邏輯斯迴歸: 參數估計	301
10.12	邏輯斯迴歸: 模型比較	302
10.13	骨質疏鬆臨床試驗邏輯斯迴歸: 配適度檢定	303
10.14	骨質疏鬆臨床試驗邏輯斯迴歸: 參數估計	304
10.15	骨質疏鬆臨床試驗邏輯斯迴歸: 新骨折機率預測	306
10.16	噪音, 鹽份攝取與高血壓: 變數	307
10.17	噪音, 鹽份攝取與高血壓邏輯斯迴歸: 參數估計	307
10.18	急性中耳炎臨床試驗: 變數說明	318
10.19	口服避孕藥與卵巢癌	319
10.20	食道癌研究: 變數名稱說明	320

10.21	DIG 臨床試驗: 變數說明 I	322
10.22	DIG 臨床試驗: 變數說明 II	323
11.1	殺蟲劑的預防蟲害劑量: 資料	348
11.2	殺蟲劑的預防蟲害劑量: 參數估計	350
11.3	殺蟲劑的預防蟲害劑量: M_2 與 M_3 參數估計	352
11.4	殺蟲劑的預防蟲害劑量: M_4 參數估計	353
11.5	殺蟲劑的預防蟲害劑量: M_6 參數估計	355
11.6	小鼠骨化畸胎試驗: 資料	358
11.7	小鼠骨化畸胎試驗: 參數估計 (標準誤)	359
11.8	膝關節鈍傷研究: 變數說明	360
11.9	冠狀動脈疾病回溯性研究: 變數說明	361
11.10	心血管心臟疾病診斷資料: 變數說明	362
12.1	子宮內膜癌: 變數說明	370
12.2	子宮內膜癌: 條件邏輯斯迴歸	371
12.3	子宮內膜癌: 一般邏輯斯迴歸	372
12.4	低新生兒體重研究: 條件邏輯斯迴歸	374
12.5	異位性皮膚炎臨床試驗: 變數說明	375
12.6	心肌梗塞配對研究: 變數說明	376
13.1	PHS 臨床試驗: 阿斯匹靈與心肌梗塞	381
13.2	PHS 臨床試驗: 阿斯匹靈與心肌梗塞參數估計	382
13.3	年齡, 抽菸與呼吸功能測試	383
13.4	年齡, 抽菸與呼吸功能測試: 參數估計	384
13.5	門診醫療系統選擇: 變數說明	386
13.6	門診醫療系統選擇: 部分資料	386

13.7	門診醫療系統選擇: 參數估計	387
13.8	統計軟體: 比例勝算配適參數形式	392
13.9	關節炎止痛藥物臨床試驗	393
13.10	關節炎止痛藥物臨床試驗: 參數估計	394
13.11	關節炎止痛藥物臨床試驗: 參數估計比較	398
13.12	子宮內膜癌研究: 變數說明	399
13.13	小細胞肺癌臨床試驗	401
13.14	小細胞肺癌臨床試驗: 變數名稱與定義	401
14.1	飲食攝取糖量與糖尿病	410
14.2	飲食攝取糖量與糖尿病: 簡單布瓦松迴歸	410
14.3	糖尿病治療臨床試驗不良反應次數: 部分資料	417
14.4	糖尿病治療臨床試驗不良反應次數: 變數說明	417
14.5	糖尿病治療臨床試驗不良反應次數: 參數估計	418
14.6	兒童下呼吸道感染: 變數說明	419
14.7	兒童下呼吸道感染: M_1 檢視個別解釋變數	420
14.8	兒童下呼吸道感染: M_1 與 M_2 巢狀模型檢定	420
14.9	兒童下呼吸道感染: M_1 布瓦松迴歸參數估計	420
14.10	兒童下呼吸道感染: M_2 布瓦松迴歸參數估計	421
14.11	吸菸與心血管疾病死亡: 世代研究	421
14.12	瓦松迴歸迴歸: 模型比較	422
14.13	汽車保險求償資料: 變數說明	425
14.14	汽車保險求償資料	425
14.15	汽車保險求償資料: 參數估計	428
14.16	兒童下呼吸道感染: M_3 調整 overdispersion 的參數估計	430
14.17	兒童下呼吸道感染: M_3 檢視個別解釋變數	431

14.18	兒童下呼吸道感染: M_4 負二項迴歸的參數估計	434
14.19	兒童下呼吸道感染: M_4 檢視個別解釋變數	434
14.20	醫療支出抽樣調: 變數說明	439
14.21	醫療支出抽樣調查: 參數估計	441
14.22	醫療支出抽樣調查: 模型比較	442
14.23	醫療支出抽樣調查: 估計 0 值數目	442
14.24	NHS 世代研究: 變數說明	443
14.25	急診室醫療糾紛次數: 變數說明	444
14.26	鎳提煉廠工人世代研究: 變數說明	446
14.27	鎳提煉廠工人世代研究: 部分個體研究資料	446
14.28	佛明罕世代研究: 變數說明	447
15.1	關節炎臨床試驗	451
15.2	抗癲癇藥物臨床試驗: 變數說明	452
15.3	群聚共變數同義詞	453
15.4	工作相關係數矩陣 $R_i(\alpha)$	460
15.5	甲癬臨床試驗: 變數說明	462
15.6	甲癬臨床試驗: 嚴重感染分率	462
15.7	甲癬臨床試驗: 邊際模型使用 GEE 估計	464
15.8	甲癬臨床試驗: 個體特定模型	467
15.9	高血壓與手腳冰冷臨床試驗	470
15.10	預防皮膚癌臨床試驗	471
16.1	三維列聯表參數個數	487
16.2	三維表中對數線性模型的符號	490
16.3	三維列聯表: 對數線性模型與邏輯斯迴歸模型	490
16.4	新生兒神經管缺陷: 變數說明	491

16.5	新生兒神經管缺陷: 資料	491
16.6	新生兒神經管缺陷: 二維列聯表	492
16.7	新生兒神經管缺陷二維列聯表: 模型比較	492
16.8	新生兒神經管缺陷二維列聯表: 對數線性模型	493
16.9	新生兒神經管缺陷: 三維列聯表	494
16.10	新生兒神經管缺陷三維列聯表: 對數線性模型	494
16.11	骨質疏鬆臨床試驗	496
17.1	方正列聯表: 模型比較	506
17.2	婚前性行為與婚外情意見表	511
17.3	婚前性行為與婚外情意見一覽表	512

圖目錄

1.1	常見醫學與流行病學的基本研究設計	8
1.2	固定世代 (fixed cohort) 長期追蹤研究, ● 起始追蹤, × 事件發生時間 點, ○ 追蹤結束時間點事件尚未發生.	12
1.3	動態世代 (dynamic cohort) 長期追蹤研究, ● 起始追蹤, × 事件發生 時間點, ○ 追蹤結束時間點事件尚未發生.	12
1.4	三種常見世代研究. (Grimes 與 Schulz, Lancet, 2002)	14
1.5	醫學實驗性研究: 臨床試驗	15
1.6	離散均一分配機率密度函數	18
1.7	伯努利分配機率密度函數	19
1.8	二項分配分配機率密度函數	21
1.9	幾何分配分配機率密度函數	23
1.10	超幾何分配分配機率密度函數	25
1.11	布瓦松分配分配機率密度函數	26
1.12	負二項分配分配機率密度函數	31
2.1	發生速率 與 人-年 (Person-Year): ● 初始追蹤, × 死亡, ○ 存活.	38
3.1	事件 (疾病) 勝算比 (OR) 與 相對風險 (RR) 的差異	97

3.2	事件 (疾病) 勝算比 (OR) 與風險比 (RR): 當事件 (疾病) 發生率小於 0.1, OR 才會是 RR 的近似值.	97
4.1	性別, 心電圖與冠狀動脈疾病: 整體與部份勝算比與信賴區間.	154
6.1	接受器操作特性曲線 (ROC curve)	192
6.2	血中雌二醇濃度與乳癌: ROC Curve	197
6.3	血中雌二醇濃度與乳癌: 1-PPV versus NPV Curve	198
10.1	樣本分率, 線性迴歸與邏輯斯曲線	274
10.2	邏輯斯曲線	275
10.3	邏輯斯曲線: 斜率變化	279
10.4	聚集資料與經驗對數勝算比	281
10.5	聚集資料與樣本分率	281
10.6	噪音, 鹽份攝取: 高血壓預測機率	308
10.7	偏誤-變異數的權衡交易 (Bias-Variance Trade-Off)	313
10.8	模型複雜程度相對於訓練資料與測試資料的預測誤差.	313
11.1	常見的殘差圖	335
11.2	殺蟲劑的預防蟲害劑量: 經驗對數勝算比	349
11.3	殺蟲劑的預防蟲害劑量: 局部殘差 u_i	351
11.4	M_2 模型: 皮爾森殘差與槓桿值	352
11.5	M_4 模型: 皮爾森殘差與槓桿值	353
11.6	M_4 模型: 模型診斷	354
14.1	簡單布瓦松迴歸模型: 事件發生速率	409
14.2	吸菸與心血管疾病死亡	422

15.1	比較邊際模型與個體特定模型差異: 黑色粗線為一條邊際模型曲線, 灰色細線為多條個體特定模型曲線.....	454
15.2	甲癬臨床試驗: 嚴重感染分率	463
15.3	個體特定模型積分後的邊際模型曲線: 灰色細線為個體特定模型的機率, 黑色粗線為個體特定模型的代表機率 ($b_i = 0$), 灰色虛線為積分後的邊際模型機率.	469
