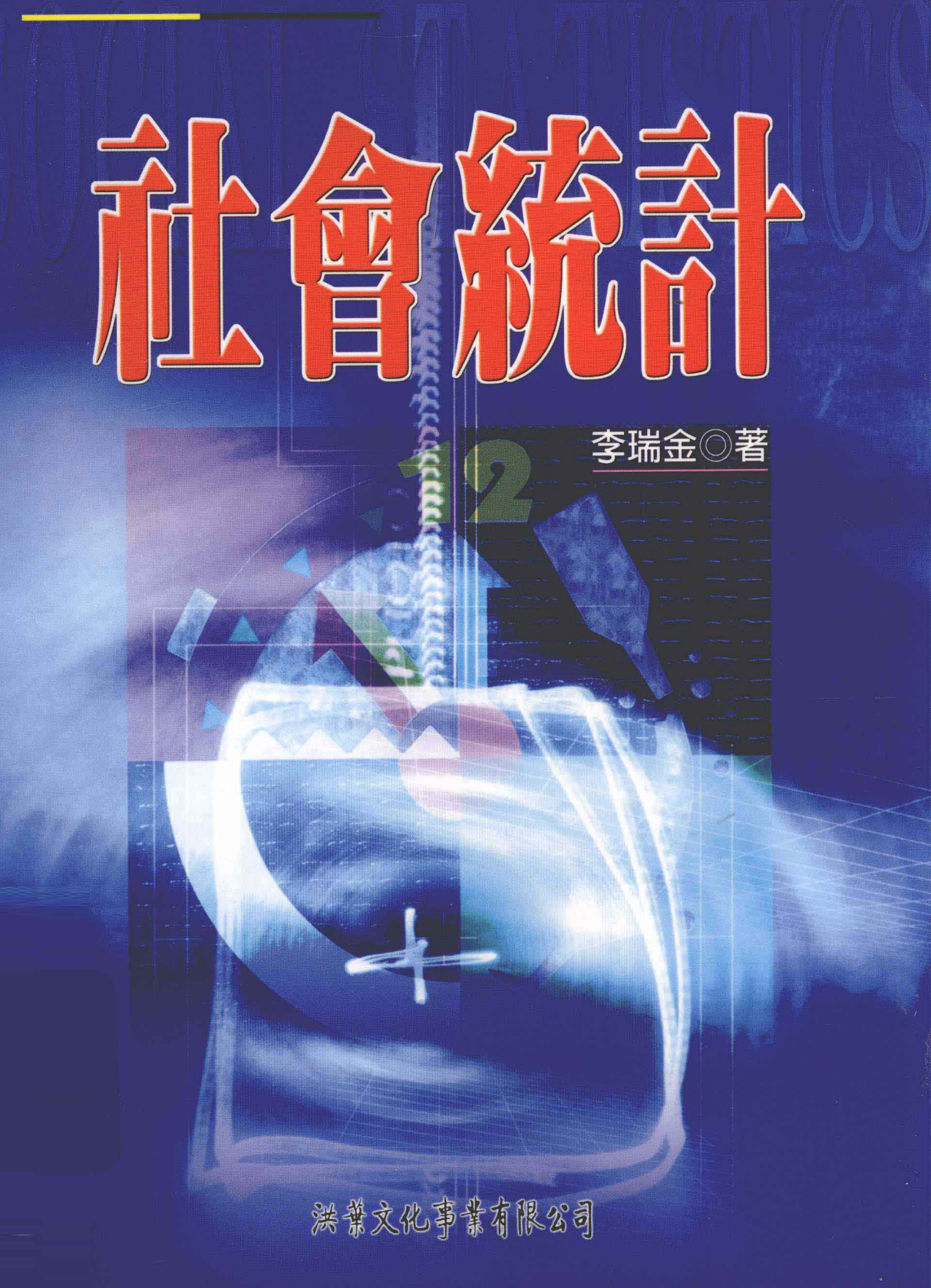


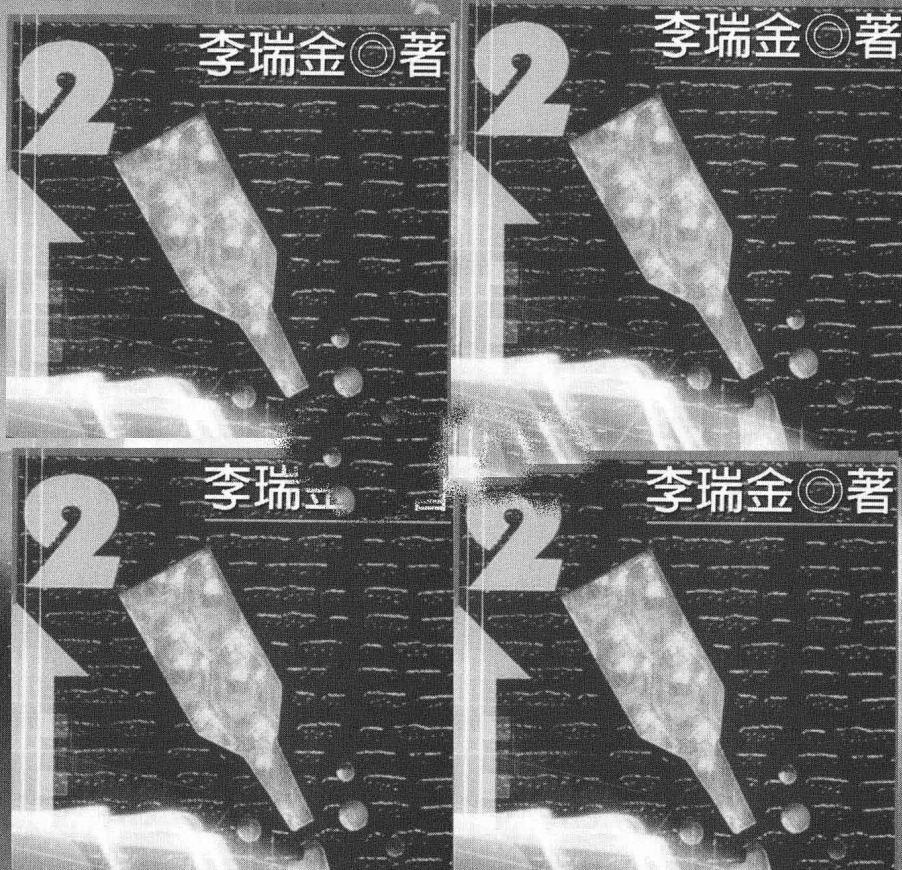
社會統計



李瑞金◎著

洪葉文化事業有限公司

社會統計



洪葉文化事業有限公司

國家圖書館出版品預行編目資料

社會統計學 李瑞金著. — 初版. — 臺北市

洪葉文化, 2004[民 93]

面；公分

ISBN 986-7553-30-6 (平裝)

1. 社會科學 — 統計方法

501.28

93015825

社會統計

作者／李瑞金

企劃主編／郭淑玲

主編／金美香

執行編輯／連茹萍

發行人／薛慶意

發行所／洪葉文化事業有限公司

登記號：局版北市業字第 5509 號

地址：106 臺北市羅斯福路三段 283 巷 14 弄 22 號 3 樓

電話：(886-2) 2363-2866

傳真：(886-2) 2363-2274

劃撥：1630104-7 洪有道 帳戶

網址：www.hungyeh.com.tw

e-mail：service@hungyeh.com.tw

門市部／電話：(886-2) 2736-2544

排版所／全凱數位資訊有限公司

版次／2004 年 9 月 初版一刷

ISBN／

定價 500 元

◎如有缺頁、破損、裝幀錯誤請寄回更換

◀ 版權所有・請勿翻印 ▶

序

統計學實際上應該稱為「數據科學」，是一門蒐集、整理、分析、顯示、解釋資料，然後再與原始問題結合，解決問題的學問。統計學在現代社會中應用相當普及，舉凡醫院實習新藥之藥效檢定，改善工業產品品質、政策評估、社會科學、農業生產等都要用到統計學。近年來由於社會問題日趨嚴重，社會學研究快速成長，社會現象經常發生變化，各種影響政策的事項牽連廣泛，因果關係錯綜複雜，使得統計技術在社會科學的應用非常廣泛。

本書力求可讀性、實用性之提高，儘量減少統計理論之研討及公式之演算，書中有關理論之探討，其所強調是概念而非數學，務使其適合社會科學系所學生之程度，引用為學生所感興趣之社會問題或社會現象之題材為例，以吸引同學對統計學的興趣，本書強調者是如何針對研究情境選用最適當統計量來分析。

本書共十五章，包括敘述統計及推論統計兩部分，前五章為統計敘述，自第六章以後為統計推論，包括無母數檢定、時間數列及多變量統計分析。本書特色是範例多，各章之後均附有習題，附錄中所載各表（特別是有關非母樹脂統計程序）亦相當完備。學生在實例中了解各統計量如何運用在社會現象中，但其基本原理及處理過程與一般統計並無差異，書中介紹方法仍可適用於其他科學，如政治、公共行政、人口、衛生、教育、生物等之統計教學，學生亦可自習題演算過程中增強其運算之技術，對同學參加研究所、高普考試及其他相關考試者，亦或有所助益，並在習得統計知識與方法後更可做為規劃學術研究及資料分析之用。

作者積多年教學及實證研究經驗，期望藉此書籍出版給予社會工作學

系學生應用「統計」的基本概念，做為資料分析及閱讀他人學術報告能力之用，社會統計為實用性之書籍，期望本書有高度之實用價值。

本書得已完成順利出版，應感謝中央大學統計研究所博士生林哲揚在推論統計部分提供許多寶貴意見，並協助全書校對，以及究助理李欣雲幫忙打字、校對，並此致謝。

李瑞金 謹識

2004 年 8 月于靜宜大學研究室

目 錄

序

第一章 緒論（表列與圖示）	1
1.1 統計學	1
1.2 社會統計學的意義	1
1.3 統計表徵	2
1.4 測量水平	4
1.5 統計數列	5
1.6 統計程序	6
1.7 次數分配	7
1.8 統計表	11
1.9 統計圖	15
習題	30
第二章 集中趨勢量數	35
2.1 集中趨勢量數之意義與種類	35
2.2 算術平均數	36
2.3 中位數與其他割值	43
2.4 眾數	52
2.5 幾何平均數	55
2.6 去頭去尾平均數	60
2.7 各種集中趨勢量數間之關係	61
習題	62
第三章 離中趨勢量數	67
3.1 離中趨勢量數之意義與種類	67
3.2 全距和四分位距	68
3.3 平均差	71

3.4	標準差	75
3.5	各種離中趨勢量數之比較	81
3.6	相對離差	82
	習題	85
第四章	偏態與峰度	91
4.1	偏態	91
4.2	峰度	97
4.3	動差	99
	習題	103
第五章	比、比例、百分比、率	107
5.1	比	107
5.2	比例	109
5.3	百分比	111
5.4	率	112
	習題	117
第六章	理論次數分配與統計假設	119
6.1	機率	119
6.2	超幾何分配與二項分配	125
6.3	常態分配	130
6.4	理論與假設	136
6.5	參數與統計常數	139
6.6	假設檢定	140
6.7	統計檢定步驟	141
	習題	147
第七章	單樣本檢定	159
7.1	樣本平均數分配	159
7.2	樣本變異數分配	163
7.3	母體平均數之檢定	165
7.4	t 分配	167

7.5	母體比率之檢定	170
7.6	常態母體變異數之檢定	171
	習題	173
第八章	兩樣本檢定	181
8.1	樣本均數差的性質	181
8.2	兩母體均數差之檢定	185
8.3	兩母體比率差之檢定	191
8.4	兩母體變異數差之檢定	192
	習題	195
第九章	估計	205
9.1	統計推定	205
9.2	點估計	205
9.3	區間估計	206
9.4	單一常態母體平均數之區間估計	207
9.5	兩獨立母體均數差 ($\mu_1 - \mu_2$) 之信賴區間	209
9.6	母體比率之信賴區間	213
9.7	常態母體變異數之信賴區間	216
9.8	兩常態母體變異數比之信賴區間	217
9.9	確定樣本大小	219
	習題	221
第十章	非參數檢定	227
10.1	單一樣本之母體中位數檢定	227
10.2	兩獨立樣本之母體中位數檢定	232
10.3	配對樣本之母體中位數檢定	235
10.4	連檢定	241
10.5	Kolmogorov-Smirnov 適合度檢定	244
	習題	246
第十一章	卡方檢定與列聯表	251
11.1	適合度檢定	251

11.2 獨立性檢定	255
習題	259
第十二章 相關與迴歸	267
12.1 相關的性質	267
12.2 相關係數	270
12.3 直線迴歸	283
12.4 估計標準誤	288
12.5 曲線相關	297
12.6 相關比	306
12.7 多元相關	313
習題	319
第十三章 相關推論	325
13.1 母體相關係數 ρ 之檢定	325
13.2 母體相關係數 ρ 之推定	326
13.3 兩相關係數差之檢定	327
13.4 相關係數之變異數分析	328
13.5 非直線相關之檢定	330
13.6 複相關複數之檢定和推定	333
13.7 母體偏相關係數之檢定和推定	334
13.8 迴歸直線	335
13.9 母體迴歸係數之檢定	338
13.10 母體迴歸係數之推定	340
13.11 兩母體迴歸係數之檢定	341
13.12 有關迴歸線之可信帶和信賴範圍	342
習題	347
第十四章 時間數列	353
14.1 時間數列的性質	353
14.2 原始資料之整理	354
14.3 直線性長期趨勢	356

14.4 非直線趨勢	363
習題	365
第十五章 實驗設計與變異數分析	367
15.1 變異數分析的意義	367
15.2 一因子完全隨機設計	369
15.3 一因子隨機化區集設計	373
15.4 二因子因子設計	376
15.5 Kruskal-Wallis 檢定	382
習題	384
附錄一 二十四個希臘字母	391
附錄二 計算用表	392
附表 1 常態曲線座標表	392
附表 2 常態曲線兩尾面積表	393
附表 3 隨機號碼表	394
附表 4 標準常態分配	400
附表 5 t 分配	402
附表 6 χ^2 分配	403
附表 7 F 分配	404
附表 8 二項分配	412
附表 9 Wilcoxon 順位和檢定的臨界值 $w(n'; \alpha)$	419
附表 10 Mann Whitney Wilcoxon 檢定	420
附表 11 Mann Whitney Wilcoxon 檢定	425
附表 12 Kruskal Wallis 檢定統計量小樣本時尾點 $w(\alpha)$	426
附錄三 名詞解釋	427
李瑞金重要重作	445

第一章 緒論（表列與圖示）

1.1 統計學

統計學（Statistics）是一種科學，它搜集、分析並解釋關於個體聚合的數字資料（Numerical data relating to an aggregate of individuals; the science of collecting, analysing and interpreting such data）。

陳超塵謂：「統計學為一種能在不確定情況下作成決策的科學方法」。

田克明教授說：「統計資料為由點計或度量社會現象或自然現象之群體的特性而得之資料。……統計方法為搜集、整理與分析統計資料並解釋統計結果之科學方法。簡言之，統計方法即搜集與處理統計資料之方法。……統計學為闡明統計方法之科學。……可分為純理統計學（Pure statistics）與應用統計學（Applied statistics）兩類。前者亦稱為數理統計學，係用數學方法闡述統計方法之原理，證明各種公式之來源。應用統計學，闡明統計方法在各種科學研究上或事業管理上之如何應用」。

統計學和數學、論理學同樣為抽象科學，和物質科學、生物科學、社會科學等具體科學相對待，然而各種具體科學都以抽象科學作基礎。統計學是一種科學方法，所有科學都須要此種方法。

1.2 社會統計學的意義

社會學辭典把社會統計（Social statistics）解釋為「以數學形式搜集、分析與組織而成的有關人類社會以及社會關係的數量資料」（Quantitative data relating to human society and social relations, collected, analyzed and organized in mathematical form）。

社會統計學，是搜集、整理、比較、分析有關社會或社會問題的數量資料的科學方法。所以，它是應用統計學。狹義言之，它是社會學方法之一種——統計法。從事社會研究工作者無不使用此種方法。近年發展之計量社會學（Quantitative sociology），即以社會統計學為骨幹。搜集社會資料（The data of social facts），整理、比較、分析之，以解釋社會現象，並歸納成社會通則。或者對社會現象擬定假設、實驗設計、分析因素並檢定假設，而成原理原則。如此種種的研究過程，非使用社會統計方法均不能達成。

1.3 統計表徵（Statistical Characteristics）

所謂表徵（The characteristic），即是關於個體、團體或文化之一種特質，和其他特質儼然有區別（That trait, property, or attribute of an individual, group, or culture which distinguishes it from others）。表現為一個單位（a single unit）或一數列單位（a series of units），因此統計單位之性能曰統計表徵。

表徵種類（Types of characteristics）：

（A）品質（Attributes）：稱為非數量表徵（Nonquantitative characteristics）或屬質的表徵（Qualitative characteristics）。其性質為非此即彼，是或否，出現或不出現。有時稱作第一類型資料（Type I data），又分下列兩種：

（1）兩分法（A dichotomy）：對生表徵（Dichotomous characteristics）係將一群人分作兩類，如男和女、瞎子和非瞎子、有電視者和無電視者等。

A 非 A		就業	無業	計	
n_A	$n_{\bar{A}}$	n	40,877	3,002	43,879 (千人)

（2）多項分類（A manifold classification or a polytomy）：將份子分成三類或三類以上，如職業分類、區域非類等。

A_1	A_2	A_3		A_R	
n_{A1}	n_{A2}	n_{A3}		n_{AR}	n

表 1.1 美國某年每十萬人殺傷和劇烈攻擊人數

地理區域	殺傷率	劇烈攻擊率
East South Central	12.45	102.6
South Atlantic	10.30	214.9
West South Central	9.11	75.9
East North Central	4.33	66.4
Pacific	3.21	49.7
West North Central	3.05	52.3
Mountain	2.76	36.3
Middle Atlantic	2.56	34.9
New England	1.26	12.1

(B) 順序或虛擬變數 (Order or dummy-variable)：可以排列成順序，不能精確地定以數字，此種資料稱為第二類型資料 (Type II data)。如：

住宅狀況： 好 可 壞
 社會身分：白領階級 藍領階級 農夫
 當然可歸納成一種模式： 高 A 中 A 低 A

n_h	n_m	n_l	n
-------	-------	-------	-----

常常以分數呈現，如社會態度：

完全同意 同意 無意見 不同意 完全不同意
 4 3 2 1 0

(C) 變數 (Variables)：稱為數量表徵 (Quantitative characteristics)，如年齡、收入、所生子女數等都可以數字表示出來，均有尺度可以測量，故稱為可測量的資料 (Measurement data)，或稱第三類型資料 (Type III data)。

(1) 不連續變數 (Discrete variables)：如家庭大小，或為 2、

為 3、為 4 等，但不能為 3.5。如所生子女數為 1、為 2、為 10 等，但不能為 1.4。

(2)連續變數 (Continuous variables)：依賴於測量之精確程度，如年齡、重量、高度、時間等可出現的任何實數。

1.4 測量水平 (Levels of Measurement)

從理論概念到實際概念之過程，無疑地非常直接。連接兩種概念須要一些決策 (Decisions)，否則彼此拉不上關係，二者無法湊在一起，達不到概括的目的。同樣的，從選擇適當的數理模型或統計模型到應用既定的研究技術或操作程序，其過程也包括許多決策。當社會現象被測量，數字體系之選擇，便是例行之事。根據尺度的性質，可將測量分為四個水平：

(一)名目尺度 (Nominal scales) 又稱類名尺度、分類尺度，在任何科學當中，簡單而基本的工作，便是分類。根據某種表徵分類，決定何者最近似，何者最有差別。譬如根據宗教分類成美以美會、長老會、天主教等，觀察宗教是否和偏見或政治保守性有關，結果發現長老教派較天主教派更具保守性。名目尺度要求各類別間有周延和互斥關係，不需有先後順序關係，各類別都有數字與之互相對應，但這些數字大多不能運算。

類名尺度為最簡單的測量標準，為使用統計程序時所必須。它具有對稱性和轉移性 (The properties of symmetry and transitivity)。關於對稱性，A 與 B 間之關係也就是 B 與 A 間之關係。關於轉移性，意即若 $A=B$ 和 $B=C$ ，於是 $A=C$ 。合而言之，若 A 階層同於 B 階層，則 B 階層同於 A 階層，若 A 與 B 為相同階層以及 B 與 C 為相同階層，於是 A 與 C 必須為相同階層。常見的名目尺度，有性別、婚姻狀況、宗教類型、血型等。

(二)順序尺度 (Ordinal scales) 某類具有某些表徵，而又不能確定具

有多少，但和他類在程度上有差別，便可據此排列順序。因此有一種常模（**A continuum**），個體或事象始得依順序排列。或者極其精確地評定等級，以致於沒有兩個個體或兩種事象同在常模之一點上。然而，通常有許多結（**Ties**），即將此歸入一單類中（**a single category**）。例如：我們說這班學生比那班學生有較高的成績；根據社會經濟地位區分家庭為「上、中上、中下和下」，或者僅分成上下兩類。能說 $A > B$ ，卻不能同時說 $B > A$ ，所以順序尺度不具對稱性。然而它具有轉移性，因此可以說：「若 $A > B$ 和 $B > C$ ，則 $A > C$ 」。

測量之順序水平（**an ordinal level**）並未表示因素間相差之大小，僅知 $A > B$ ，但不知大多少。亦未顯示： A 和 B 之差小於 C 和 D 之差。常見的順序尺度有出生別、社經地位、滿意程度等。

（三）**間距尺度（Interval scale）**測量不僅評定個體所具有某種表徵成份之等級，而且也顯示個體間之正確距離。測量之間距尺度水平需要建立測量單位（**a physical unit of measurement**），大家都贊成的共同標準，雖經迭次使用，仍有共同的結果，如長度用公尺或英尺，重量用磅或公克，時間用秒，溫熱用度，收入用元等。唯間距尺度的數字間可以用減法運算，但不能做除法運算，如智商 140 的人不能說是智商 70 的人的兩倍。

（四）**比例尺度（Ratio scale）**一個分數兩倍於另一個分數，或二分之一倍於第三個分數，比例尺度所含各類別間不但大小之間有單位多少可言，且有絕對零點（**absolute zero**）為起點，因此比例尺度可以作除法運算，常見的比例尺度有體重、收入人數等。這種比例尺度常被認為是測量之頗高標準。

1.5 統計數列（Statistical Series）

分析數字資料須依序排列，而此種依照順序排列的數字資料便稱為

數列或分配 (distribution)。可分下列四種：

- (一) 時間數列 (Time series) 資料之排列和分組，依發現之時間先後所得之數列，稱為時間數列。
- (二) 空間數列 (Spatial series) 資料之排列和分組，按地理之位置所得之數列，稱為空間數列或空間分配。
- (三) 品質數列 (Series of attributes) 或稱為屬性數列，係按不同品質或同一品質的特性之差別程度而分類編排的統計資料。
- (四) 次數數列 (Frequency series) 又稱次數分配或變量數列 (series of variates)，係依照數量大小順序分組排列之數字資料。可分為連續數列 (Continuous series) 和非連續數列 (Discrete or discontinuous series)。

1.6 統計程序 (The Statistical Procedure)

搜集原始資料 (Primary data)，有下列幾種步驟：

- (一) 計畫一種研究 (Planning the study)：在自己的研究領域中，尋找所要研究的種種問題。按照目前的環境，在人力財力許可下選擇並確定一個迫切而適當的研究題目。進而從報紙、雜誌、書籍、報告等搜集次級資料 (Secondary data)，並分析比較之。
- (二) 設計問項 (Devising the questions)、編製調查表 (Making the schedule)：根據研究題目，劃定研究範圍，設計一些必不可缺少的問項。這些問項要清晰 (Clarity)。所希望的回答要具體 (Objective)、要能列表。
- (三) 選擇樣本類型 (Selecting the type of sample)：假若不是全體性地搜集資料，便需要樣本。隨機樣本 (The random sample)、系統樣本 (The systematic sample)、叢集樣本 (The cluster sam-

ple）、分層樣本（The stratified sample）、立意樣本（The purposive sample）、連續抽樣（Sequential sampling）等，究竟以何者為宜？

- （四）**收集事實**（Using the schedules to obtain the information）：用郵寄問卷表法（The questionnaire）或訪問法（The interview）？或兩者並用？
- （五）**整理調查表**（Editing the schedules）：需要點計（Computing）、編號（Coding）、解釋（Deciphering）、校核（Checking），並檢查是否完善（Examining for completeness）。
- （六）**整理資料**（Organizing the data）：現今是電腦時代，大規模的數量資料可由電腦處理，使研究方法邁入另一境界。不過，電腦處理數量資料之方法，尚未普遍應用，傳統的整理資料之方法仍未失其重要性，尤其在個人從事小研究時。歸類、化簡、計算代表數值等，都儘可能地使用工具。
- （七）**製表、繪圖**（Making finished tables and charts）：統計表更具執簡馭繁的功效。統計圖更使人一新耳目，更易使人有一種鮮明的輪廓。
- （八）**分析結果**（Analyzing the findings）：分析為統計之最重要工作。有了此項工作才能使社會事實印證原理原則。

以上可以說是研究步驟，前半屬於調查階段，後半屬於統計階段。所以說：社會調查和社會統計是一體之兩面。廣義的統計程序，可以說是研究的程序；狹義的統計程序只是整理、分析資料而已。

1.7 次數分配（Frequency Distributions）

- （一）**整理資料**（Grouping the data）下列資料為 93 地區投票百分比之