科學圖書大庫

大專用書

統計學

譯者 林元興 校閱 鄭堯柈

徐氏基金會出版

科學圖書大庫

大專用書

統 計 學

譯者 林元興　校閱 鄭堯柈

徐氏基金會出版

徐氏基金會科學圖書編譯委員會
監修人　徐銘信　發行人　王洪鎧

科學圖書大庫

中華民國六十八年三月十六日六版

統　計　學

基本定價　3.00

譯者　林元興　國立政治大學副教授
校閱　鄭堯柈　國家科學委員會研究教授

本書如發現裝訂錯誤或缺頁情形時，敬請「刷掛」寄回調換。謝謝惠顧。

(67)局版臺業字第1810號

出版者　財團法人臺北市徐氏基金會　臺北市郵政信箱53-2號　電話7813686號 7815250號
發行者　財團法人臺北市徐氏基金會　郵政劃撥賬戶第15795號
承印者　大興圖書印製有限公司 三重市三和路四段一五一號　電話9719739

目 錄

第一章　緒論

「統計學」(Statistics) 原係國家搜集人口與經濟資料的技術。嗣後却逐漸演變爲一種科學的分析方法，現已應用在各種社會科學與自然科學中，且形成數學的一種主要分支。現代統計學的目標與方法，可藉以下的實例予以說明。

1-1 實例

在外國每逢選舉總統前，各民意測驗機關都想推測誰會得勝；尤其是要預計各候選人的得票比例。如對全部選民一一普查，工作實過份繁重。唯一可行的辦法是調查少數的選民，期能由該等樣本（ sample ）中，推定群體（ population ）全部的資料，此即「統計推論」(Statistical inference）的典型實例：由樣本的（投票）比例以推論群體的（投票）比例。

各民意測驗機關咸認，這是一項冒險的工作。群體的正確投票比例，須俟選舉日所有選票均點計後方能揭曉。惟選取樣本（ sampling ）的方法如安排恰當，則樣本比例必與群體比例相近。故可由觀察的樣本比例 (P)，以推定未知的群體比例 π，其方式如下：

$$\pi = P \pm \text{誤差} \tag{1-1}$$

惟現有若干問題發生，譬如「誤差究有多大？」與「推測的正確性如何？」

因爲這個實例正是本書的精髓部份，故須引用第七章的用語嚴加說明（讀者嗣後自會獲得證明與全盤瞭解）。

選取樣本如按隨機（ random ）方式，且數目充份，則下列方程式必達 95 %的信賴程度 (confindence)，即

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

其中 π 與 P 分別代表群體比例（ population proportion ）與樣本比例（

Sample proportion)，而 n 代表樣本數目(Sample size).

茲舉一實例以說明該公式之應用，設已選出 1,000 位選民，其中 600 位擬投票予甲黨推出之候選人。將 0.60 的樣本比例代入公式 (1-2)，即得

$$\pi = .60 \pm 1.96\sqrt{\frac{.60(1-.60)}{1000}}$$

演算後約為

$$\pi = .60 \pm .03 \tag{1-3}$$

故在 95% 的信賴程度下，投票予甲黨的群體比例應在 0.57 至 0.63 之間。

此即「信賴區間」(confidence interval)，而信賴區間之推定，係本書的一個主要研討題目。另外一個題目係「假設檢定」(test hypotheses). 舉例而言，得檢定乙黨候選人必贏得該次選舉之假設，惟根據方程式 (1-3) 的資料，一定會放棄該等假設，蓋由樣本調查的結果，甲黨占 57% 至 63%，實居大多數，自應放棄乙黨得勝的假設。一般而言，信賴區間與假設檢定十分相近；以後由許多實例中自可看出其求法亦相同。

接着要對方程式 (1-3) 提出若干重要的說明：

1. 推定的結果並非完全正確；蓋只有 95% 的信賴程度。故有時也會發生錯誤——此即運氣不佳，抽出一組不妥的樣本所致。即使群體中贊成甲黨的人數不及半數，惟可能將大多數的甲黨人士選為樣本，這種情形雖不常見，但一旦產生後，根據 (1-3) 所求的結論必大錯特錯。運氣不佳機會雖鮮，惟並非沒有發生的可能，故推定的結論僅有 95% 的信賴程度。

2. 樣本的數目如予增加，運氣的成份必隨之減少；調查的選民愈多，由乙黨居多數的群體中，選出甲黨占多數的樣本，機會愈少。故預測的結果愈為精確。由方程式 (1-2) 中即可得證；蓋該式表示樣本愈增，則誤差愈小。現如將樣本增為 10,000 位選民，且其中贊成甲黨的比例仍為 0.60，則 95% 的信賴區間必更為精確：

$$.60 \pm .01 \tag{1-4}$$

3. 如有人認為 95% 的信賴程度不夠理想。「應將推定結論的信賴程度提高至 99%」。則有兩途可循。一是增加樣本的數目；此需增加調查費用與勞力，結果獲得之推定區間當與 (1-4) 同樣精確，惟却可提高信賴程度。如無法增加樣本的數目，則只有減少推定結論的精確性，俾提高信賴程度

——此卽，將贊成甲黨人士的比例改爲：

$$.60 \pm .02$$

預測的精確性愈低，則其信賴程度愈高。在此範圍內，得採用二種辦法，以免產生錯誤的結論。一是將推定的精確性儘量放寬，以致永遠不會發生牴觸的現象（註一）。另一個辦法是將全部群體均列爲樣本（註二）加以調查；惟此非統計學的方法——僅係一種點計的工作，不必採用。一般的統計結論均附有若干不確定的程度。

1-2 歸納法與演繹法

圖 1-1 係表示歸納法與演繹法不同之處。歸納法（Induction）是將特殊的現象轉化爲一般的現象，（如由統計學而言）卽由樣本推測群體。演繹法（Deduction）正好相反——由一般的現象轉化爲特殊的現象；卽由群體求得樣本（註三）。方程式(1-1)卽歸納法；因其係根據樣本比例以推測群體比例。惟此須事先着手演繹的問題。由方程式(1-1)中可以看出，如欲獲得歸納的結論（卽群體比例可根據樣本比例加以推測），須先根據演繹的事實（卽樣本比例可能與群體比例相近）。

本書第二章至第五章均屬演繹法。舉例而言，其中包括機率（probability），機率本身用途甚多（譬如應用在博奕理論 Game theory）；惟對本書第七章至第十章的統計歸納法貢獻更大。總之，前六章所要研究的是：「在既定的群體中，樣本的出現情形如何？樣本是否正合目標的需要？」須俟該等演繹問題解決後，方能進展至統計推論 (Statistical inference) 部份。故以後數章的研究問題轉變爲：「由觀察的樣本資料中，推測未知的群體資料，其精確性究有多大？」

1-3 選樣的原因

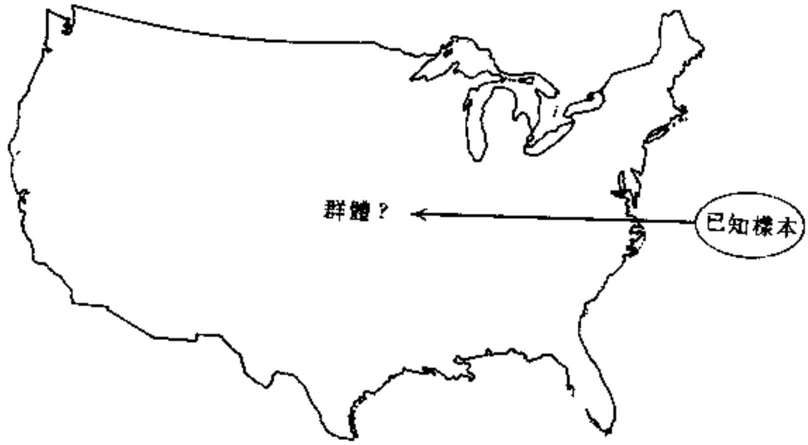
統計學中慣用選樣（譯者註：卽選取樣本之簡稱）的方法，而不研究群體全部，其原因有三：

(1) 資源匱乏（limited resources）.
(2) 資料不足（limited data available）.
(3) 檢定方法對樣本具破壞性（destructive testing）.

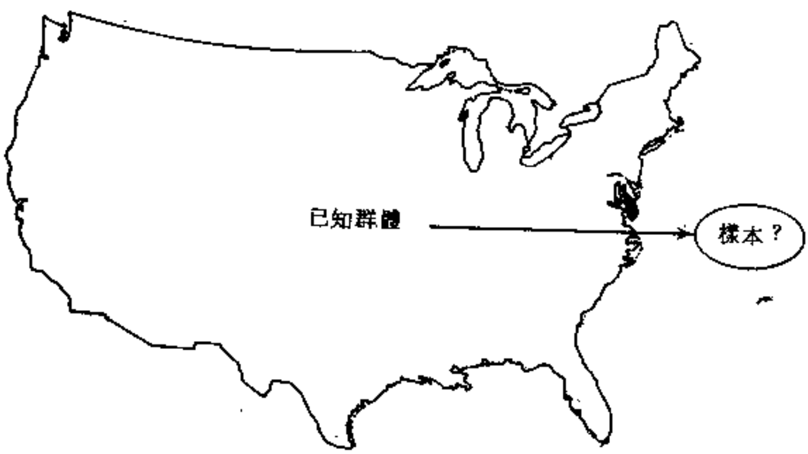
資源總是不夠使用。譬如選情預測的實例，卽無法籌措充裕的資金

，俾調查全部群體；但此非舉辦抽樣的惟一原因。

2. 有時花費大批費用，僅能獲得少數樣本。舉例而言，某人類學家欲證實甲、乙兩島的民族發展並無關連，且各具有獨特的體重、身高等徵狀。惟兩種民族無法完全加以比較。只能根據甲島現存50位居民與乙島現存100居民加以推測。故有時樣本數目早已固定，並非研究經費所能左右。



(a)



(b)

圖 1-1 歸納法與演繹法的比較 (甲) 歸納法 (統計推論) (乙) 演繹法 (機率)

工商業中亦有許多類似實例。設有一種自稱效率較高的機器，現擬依據檢定結果，決定是否添購一部。惟品質管制(quality control)人員實無法觀測該等機器生產的所有產品。僅得觀測少數樣本，再據以推測該機器效

率的高低。

3. 選樣工作可能係具破壞性的檢定。舉例而言，設某工廠現已生產一千個燈泡，且欲求其平均燃燒壽命。自然不能將燈泡全部燒壞為止。

1-4 選樣方法

統計學與工商業或其他的行業，均須避免運氣不佳(bad luck)與處理不善(bad management)的情況。舉例而言，設有某人與你打賭一百元，認為你擲一個骰子必得一點，否則包賠。若你接受，惟擲出骰子後竟得一點；故某人得勝。在此實例中，某人實處於處理不善的地位，却藉着絕佳好運以彌補其不利。惟一平反的辦法是要求繼續再玩，這次——當然要用自己的骰子。

如再觀察選情預測的實例，當可發現甲黨的樣本比例如含有以上任何一種不利情況(或兩者俱備)，即不能精確代表群體比例。選取樣本的方法無論如何妥為處理與設計，惟因運氣不佳，可能還會由乙黨居多數的群體中，抽出甲黨超過半數的樣本。方程式(1-2)正可說明這個情況；惟問題是選樣的運氣好壞，與處理方式無關。由該方程式得知，預防運氣不佳的最好辦法即「繼續再玩」；由統計學上而言，增加樣本數目，可以提高推定的可靠性。

另一問題係選樣方法處理不善或有偏誤。舉例而言，如擬選取若干選民作為樣本，自不能以電話名簿作為群體，蓋窮人無法裝設電話，即無受調查的機會。

其他尚有許多選樣偏誤的實例，俯拾皆是，頗堪玩味。譬如在街上向行人調查意見即含有偏誤，蓋樣本泰半均為溫雅有禮、衣著入時的人士；而粗魯的工人與辛勞的主婦却被忽略。民意代表也不能以所收到函件作為選民的正確意見，蓋其較為激進，其中尚包括一味孤行的人士與遊說份子。

正確選樣中最簡單的方法，係令群體中每一構成份子均有同等入選的機會。此即「隨機」(random)選樣的定義(註四)。選取樣本不能馬虎；應慎重設計。在某大都市街頭調查一千位人民的意見，即不能當作該國全體選民的隨機樣本。而應由全國東南西北普遍各選出若干。樣本如係隨機選出，方可免於偏誤(bias)，且始能符合機率論(probability theory)的要求，並按方程式(1-2)的形式，進行科學化的推論。

有的問題僅具有非隨機樣本(nonrandom sample)。雖無法引用機率論，惟仍可作為正確推測的依據——或可稱此為「推論的技藝」(art of inference)。該等技藝雖甚重要，但無法在初等統計學中研討；故本書僅

就隨機樣本進行科學方法的推論。至於選樣方法擬留待第六章再作討論。

(註一) 例如 $\pi = 0.50 \pm 0.50$

(註二) 調査的對象較全部群體略少亦無不可。蓋計算當選人之得票，無須調査全部選民，只要其中一方贏得過半數卽已足矣。(自然有些人在受調査時的意見，與在實際投票的態度並不一致，惟本書並不研究這種問題)。

(註三) 拉丁文中，字首 *in* 係代表「歸向」(into or toward)之意，若由群體出發，則歸納法(*in*duction)係歸向群體。而字首 de 係表示「離去」(away from)之意。故演繹法(*de*duction)係離去群體。總之，統計推論係根據歸納法。

(註四) 嚴格地說，此應稱爲「簡單隨機選樣」(simple random sampling)，蓋因尚有較複雜的隨機選樣。

第二章　樣本敍述統計

2-1　導言

統計學的主要目的，旨在利用樣本以推論群體。惟最初須將樣本簡化，並縮減爲若干敍述性的數字（ descriptive numbers ），此稱爲樣本「統計量」(Statistic)（註一）。

在第一章的簡單實例中，該民意測驗機關若將一千位選民意見詳爲記載，必然獲得一串像甲甲乙甲乙……的字列，其中的甲或乙分別代表贊成甲黨或乙黨的人士。惟敍述該樣本的最佳方式，莫過於利用一個簡單的統計量 P 代表，P 即贊成甲黨的樣本比例；再據以推論 π，π 係群體比例。自然，該統計量之計算甚爲簡單。在前一章的實例中，計算樣本比例 (0.60)，僅須點計贊成甲黨的人數 (600 位)，再除以樣本數目（ $n=1,000$ 名)即可求得。

現擬另舉二例，以說明統計量的計算：

(甲) 一顆骰子連擲 50 次之結果。

(乙) 二百名男子之平均身高。

2-2　次數表與圖形

(甲)　間斷變數 (discrete varible 亦可譯爲間斷變量)

如將骰子每次擲出之點數 X 詳爲記載，必得 1，2，…6，等數字。X 稱爲間斷隨機變量（discrete random varible），蓋其數值僅取有限個（或雖無限，但可數）。

一顆骰子連擲 50 次，即得 50 個數，如表 2-1 所列。

表 2-1　一顆骰子連擲 50 次之結果

6, 2, 2, 3, 5, 1, 2, 6, ……………………… 4, 2.

接着利用表2-2，以劃「正」字記號的方法，將六種點數出現結果予以簡化。第三行係表示次數（frequency）f（或稱出現次數），譬如9即表示一點的出現次數；亦表其結果爲9/50．該項比例（0.18）即稱爲相對次數（relative frequency），以符號（f/n）表示；一併計算在第四行。

表2-2　骰子出現點數之次數與相對次數

(1)	(2)	(3)	(4)
點數	正字記號	次數 (f)	相對次數 (f/n)
1	正　正	9	.18
2	正　丅	12	.24
3	正　一	6	.12
4	正　下	8	.16
5	正　正	10	.20
6	正	5	.10
		$\sum f = 50 = n$	$\sum (f/n) = 1.00$

其中Σf表「所有f的總和」

第三行的資料稱爲「次數分配」(frequency distribution)，並將其繪成圖2-1．第四行的「相對次數分配」(relative frequency distribution)亦可繪成同樣圖形；如詳爲觀察，該等圖形除縱軸的計算單位(vertical scale)有所差別外，其他一概相同。故將圖2-1的縱軸單位稍予變更，即變成相對次數分配。該圖直接顯示樣本的實驗結果。

(乙)　連續變數（continuous varible亦可譯爲連續變量）

設由某群體選出樣本二百名男子，以吋爲計算單位，將其身高逐一記載。最後目的旨在推論全部群體之平均身高；惟首先須簡化樣本資料並加敘述。

該實例中的身高（以吋計算），即隨機變數X。在此情況下，X係連續變量；蓋各人之身高可得任意值，例如63.328吋（註二）。現如觀察該特定值X，並無意義；因爲不會再出現與64.328吋相等的身高。惟可將其記入某一組(class)或區(cell)內，即於表2-3第三行58.5吋至61.5吋間劃一記號。然後再將次數與相對次數按以上方式列成一表。

「區」得隨意制定，惟一般均遵照下列原則：

1. 區數不能過於繁瑣或過於簡略。

2. 區中點(cell midpoint)應取整數，蓋嗣後須作為區內所有樣本數值之代表。

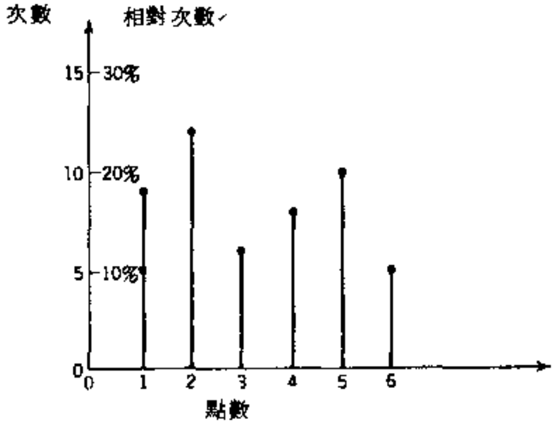


圖 2-1 骰子出現點數之次數與相對次數

表2—3 二百名男子樣本身高之次數與相對次數

區數	(1) 區界	(2) 區中點	(3) 正字記號	(4) 次數 f	(5) 相對次數 f/n
1	55.5–58.5	57	丅	2	.010
2	58.5–61.5	60	正 丅	7	.035
3	61.5–64.5	63	正 正 正 正 丅	22	.110
4	.	66	.	13	.065
5	.	69	.	44	.220
6	.	72	.	36	.180
7	.	75	.	32	.160
8	.	78	.	13	.065
9	.	81	.	21	.105
10	82.5–85.5	84	.	10	.050

$\sum f = 200 = n \quad \sum f/n = 1.00$

圖2-2 表示二百個觀察值分成各區之情形，各觀察值均用一點代表。爲簡化起見，觀察值均確實記載，而不進位(rounded off)。(舉例而言，區寬 cell width 若爲一個單位，則在分區前，須將觀察值進位至個位數 to the nearest integer).

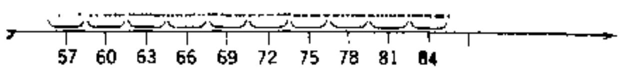


圖 2-2 觀察值之分區，即表 2-3 最初兩行之說明

接着將分區資料繪成圖2-3 。該項次數分配，或稱直方圖(histogram)，係使用直方形代表整區內觀察值之次數，而不在區中點繪直線。

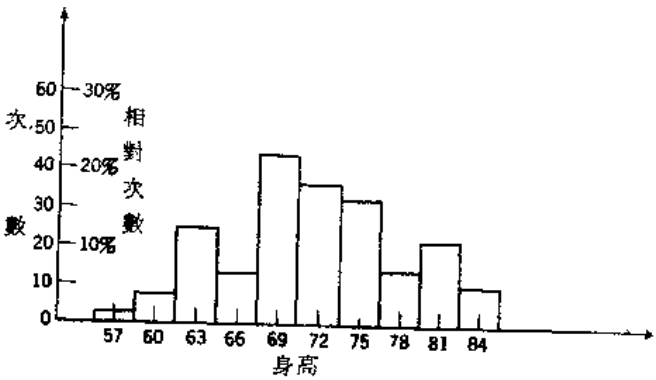


圖2-3 200 名男子身高之次數與相對次數

現擬討論如何利用一個敍述量數(descriptive measure)，或稱樣本統計量，以代表一組樣本次數。事實上，有兩種用途極廣的敍述方法：一是分配的中心點(central point)，另一係其離勢(spread)。

2-3 中心值 (位置量數 *measure of location*)

表示次數分配的「中心」(center)，方法甚多。其中之衆數(mode)、