

初等統計學

習題及複習題詳解

1985年第四版

原著者 R. J. Wonnacott T. H. Wonnacott

譯著者 張 秀 惠

曉園出版社

版權所有・翻印必究

四 版

1988年11月第一次印刷發行

初等統計學

習題及複習題詳解

定價：新臺幣 220 元 港幣 60 元

原著者：R. J. Wonnacott • T. H. Wonnacott

譯著者：張 秀 惠

發行人：黃 旭 政

發行所：曉園出版社有限公司

臺北市青田街7巷5號

電話：(02) 394-9931 (六線)

郵撥：1075734-4 號

臺大店：臺北市新生南路三段96號之三

電話：(02) 3917012 • 3947375

重南店：臺北市重慶南路一段115號

電話：(02) 3313360 • 3149580

工專店：臺北市新生南路一段6之9號

電話：(02) 3968664

逢甲店：臺中市西屯區文華路113號

電話：(04) 2512759 • 2546663

淡江店：臺北市淡水鎮英專路71號

電話：(02) 6217840

香港所：九龍又一村達之路30號地下後座

電話：3-805807 3-805705

印刷所：復大印刷廠

臺北市武成街36巷16弄15號

出版登記：局版臺業字第 1244 號

著作執照：臺內著字第 號

前 言

研習理工的同學，都有一種認識，那就是：一本書的習題往往是該書的精華所在，藉着習題的印證，才能對書中的原理原則澈底的吸收與瞭解。

有鑒於此，曉園出版社特地聘請了許多在本科上具有相當研究與成就的人士，精心出版了一系列的題解叢書，為各該科目的研習，作一番介紹與鋪路的工作。

一個問題的解答方法，常因思惟的角度而異。曉園題解叢書，毫無疑問的都是經過一番精微的思考與分析而得。其目的在提供對各該科目研讀時的參考與比較；而對於一般的自修者，則有啓發與提示的作用。希望讀者能藉着這一系列題解叢書的幫助，而在本身的學問進程上有更上層樓的成就。

初等統計學

習題及複習題詳解

(目 錄)

第一章	統計學的本質	1
第二章	敘述統計	13
第三章	機 率	37
第四章	機率分配	69
第五章	兩個隨機變數	99
第六章	抽 樣	141
第七章	點估計	167
第八章	區間估計	191
第九章	假設檢定	227
第十章	變異數分析	247
第十一章	直線配適	273
第十二章	簡單迴歸	281
第十三章	複迴歸	299
第十四章	迴歸理論延伸	323
第十五章	相 關	353
第十六章	無母數統計	383
第十七章	卡方檢定	407
第十八章	最大概似估計	425
第十九章	貝氏推論	435
第二十章	貝氏決策理論	457

第一章 統計學的本質

習 題

1-1 下表是美國最近六屆總統大選中民主黨與共和黨候選人名單，括號內的數字是選前蓋洛普民意調查對 1500 位選民所作的調查。（正如前述，我們仍然忽略美國的其他黨派，且假設該樣本為一簡單隨機樣本）。

年	民主黨	共和黨
1960	甘迺迪（51%）	尼克森（49%）
1964	詹 森（64%）	高華德（36%）
1968	韓福瑞（50%）	尼克森（50%）
1972	麥高文（38%）	尼克森（62%）
1976	卡 特（51%）	福 特（49%）
1980	卡 特（48%）	雷 根（52%）

(a) 試就每一屆大選中，建立一支持民主黨候選人之 95% 的信賴區間。

(b) 試指出下列實際結果，那些超過上題所求得之信賴區間以外：

1960	甘迺迪	50.1 %
1964	詹 森	61.3 %
1968	韓福瑞	49.7 %
1972	麥高文	38.2 %
1976	卡 特	51.1 %
1980	卡 特	44.7 %

解 (a) 1960 年： $\pi = 0.51 \pm 1.96 \sqrt{\frac{0.51 \cdot 0.49}{1500}} = 0.51 \pm 0.0253$

$$1964 \text{ 年：} \pi = 0.64 \pm 1.96 \sqrt{\frac{0.64 \cdot 0.36}{1500}} \\ = 0.64 \pm 0.0243$$

$$1968 \text{ 年：} \pi = 0.50 \pm 1.96 \sqrt{\frac{0.50 \cdot 0.50}{1500}} \\ = 0.50 \pm 0.0253$$

$$1972 \text{ 年: } \pi = 0.38 \pm 1.96 \sqrt{\frac{0.38 \cdot 0.62}{1500}} = 0.38 \pm 0.0246$$

$$1976 \text{ 年: } \pi = 0.51 \pm 1.96 \sqrt{\frac{0.51 \cdot 0.49}{1500}} = 0.51 \pm 0.0253$$

$$1980 \text{ 年: } \pi = 0.48 \pm 1.96 \sqrt{\frac{0.48 \cdot 0.52}{1500}} = 0.48 \pm 0.0253$$

(b)由於 1964 年與 1980 年的區間均未包含實際值，故此二區間錯了。

1-2 某一調頻廣播電台為提高其服務品質，特委託某機構進行一項市場調查，以 500 名聽眾為對象，調查他們對於古典音樂與流行樂曲之偏好情況。調查結果以年齡與性別區分如下表所示：

喜愛古典音樂之人數			喜愛流行樂曲之人數		
年 齡	性 別		年 齡	性 別	
	男	女		男	女
25 歲以下	19	26	25 歲以下	63	45
25-50 (歲)	38	34	25-50 (歲)	38	33
50 歲以上	48	60	50 歲以上	44	52

(a)試計算下列各問題之估計值，並分別建立 95 % 之信賴區間。

- (1)喜愛流行樂曲之年輕男性（25 歲以下）所佔之比例。
- (2)喜愛流行樂曲之年輕人（25 歲以下）所佔之比例。
- (3)喜愛流行樂曲之男性所佔的比例。
- (4)喜愛流行樂曲之女性所佔的比例。
- (5)喜愛流行樂曲之人士所佔的比例。本題與前兩題有何關係？

(b)在問題(a)中，你應該做那些假設？

解 (a)(1) $n = 500$

$$P = 63 / 500 = 0.126$$

$$\pi = 0.126 \pm 1.96 \sqrt{\frac{(0.126)(0.874)}{500}}$$

$$= 0.126 \pm 0.029$$

$$(2) P = (63 + 45) / 500$$

$$= 0.216$$

$$\pi = 0.216 \pm 0.036$$

$$(3) P = (63 + 38 + 44) / 500 = 0.29$$

$$\pi = 0.29 \pm 0.040$$

$$(4) P = (45 + 33 + 52) / 500 = 0.26$$

$$\pi = 0.26 \pm 0.038$$

$$(5) P = (63 + 38 + 44 + 45 + 33 + 52) / 500 = 0.55$$

$$\pi = 0.55 \pm 0.0436$$

(b)此 500 人之樣本為隨機樣本且 P 的分配為常態分配。

1-3 下列各項抽樣中可能產生那些誤差？如何消除？

- (a)某參議員欲從他所收到的信件中，估計出支持槍械管制法案者所佔之比例。這些信件中有 132 人表支持，429 人則持反對意見。
- (b)對於日間托嬰中心提供地區性補助的機構，欲瞭解這些當事人家長對於托嬰時間安排之看法。在所進行訪問的 2180 戶家庭中有 960 戶的回答是「上午 9 點至下午 5 點」。
- (c)某大學為估計獲得 MBA 學位者，在 10 年後之平均所得，特別在這批學生回到學校開同學會時，詢問每一位學生。其結果是：281 位十年前的畢業生中有 56 個返校參加聚會，而且其中只有 14 位願意提供有關他們所得的資料。

解 採用隨機抽樣法抽取樣本資料，則可消除可能造成的偏誤。

- (a)除了支持與反對意見外，可能還有很多中立意見。
- (b)會完全遺漏“在上午 9：00 至下午 5：00 這段時間在外工作的夫妻的意見”。
- (c)只詢問參加聚會的校友，會遺漏那些在遠地較忙碌的校友或較不願意透露所得的人。

1-4 為檢定維他命 C 預防感冒之效果，T. W. Anderson 曾經在 Toronto 進行一項研究，他將自願受測者隨機地分成兩組（每組 400 人），其中處理組的每個人均給予服用維他命 C，而控制組的每一個人則給予服用類似卻無療效的藥丸（兩種藥丸表面上看起來完全相同，毫無差別）。結果發現在冬天的時候兩組中不曾感冒的人數之比例分別為 26 % 與 18 %。（消費者報導，1976）

- (a)為什麼要給予控制組的人服用看似維他命 C 的藥丸？
- (b)如果我們假設這些自願受測的樣本是來自一龐大母體的隨機樣本，

試建立此一母體中，服用維他命丸而確能免除感冒之比例的 95% 信賴區間。

(c)就控制組而言，重複問題(b)。

(d)試對維他命 C 之效果作一簡單的結論。

解 (a)以便將其效果和服用維他命 C 的效果比較。

$$(b) \pi = 0.26 \pm 1.96 \sqrt{\frac{(0.26)(0.74)}{400}} = 0.26 \pm 0.0429$$

$$(c) \pi = 0.18 \pm 1.96 \sqrt{\frac{(0.18)(0.82)}{400}} = 0.18 \pm 0.0377$$

(d)實驗結果證實維他命 C 確有預防感冒的效果。

1-5 已知某一包含 n 種「處理」的實驗中，各種處理的效果完全未知（例如，酒精、義務教育、心理療法），你如何設計一實驗來評估這些「處理」的效果。提示，你的討論應包括：

(a)如何選擇受測者？

(b)如何分派處理組與控制組？

(c)那些人應該處於毫不知情（盲目）的情況？

(d)如何評估其效果？

解 (a)選擇實驗目的所需研究之對象（例如：研究補習對學生在校成績的影響，其研究對象即為某一年級之學生）。

(b)以隨機的方式將研究對象分成處理及控制兩組，（對處理組之學生施以課後補習）。

(c)使評估者（evaluator）與受測者本身對此實驗毫無所知。（使老師與學生不知道此實驗之進行）

(d)從學生成績來比較兩組學生成績是否有顯著差異。

1-6 假設在本節中所述對 28 位婦女進行照顧嬰兒的實驗中，每位受測者（婦女）均給予一個代號，如下所示（單身婦女以大寫字母表示）：

A	a	b	B	C	D	c	d	e	E	f	F	g	h	j
k	G	m	n	p	H	J	q	r	s	t	K	u		

將上列名單隨機指派至控制組與處理組的方式之一是，登記出每一个人的名字作成卡片，並放置在一頂帽子內，然後全部倒出來平均分為兩組。然而，有時候這些受測者的全部名單在開始時不得而知，例如，這些

人可能剛好是一個接著一個進來，一直到研究者確定所須的資料已經足夠為止。因此，我們可能須要採取另一種隨機化的方式，例如，當每一受測者進入醫院後，研究者即根據擲一枚硬幣視其出現正面或反面的方式，來指派她到底分配在控制組或處理組。

(a)對於上列28位婦女的名單中，以擲硬幣視其出現正、反面的方式決定每一位婦女應分配在那一組的隨機化是如何地操作？（或者使用本書附錄I的隨機數字表來決定，其中單數代表硬幣之正面，偶數代表反面）。然後，將出現「正面」的受測圈起來；並將之指派至處理組。

(b)根據題意，10個大寫字母代表10個單身婦女，則(a)中的隨機化過程是否能夠將剩下的婦女平均分配至處理組與控制組。

(c)假設以上的實驗重複進行5次，結果將為何？

(d)為確定你對問題之猜測是否正確，請重複此種隨機過程5次，看看結果而言，這種指派是否剛好每組人數相等？這是否足以說明隨機化將消除所有的其他影響因素？

解 (a)利用附表一之隨機號碼表：（從第6列第一行開始，橫找）：“偶數”分配至實驗組，單數則分配至控制組。

Ⓐ ① ② ③ B ④ ⑤ c ⑥ e ⑦ ⑧ F g h ⑨
k ⑩ m n p H ⑪ q r ⑫ t ⑬ ⑭

(b)單身婦女分配至實驗組者有7個，因此分配並不均勻。

(c)若重複分配多次，則可望單身婦女平均分配至實驗組者有5人。

	單身		已婚	
	實驗組	控制組	實驗組	控制組
第一次	7	3	7	11
第二次	5	5	8	10
第三次	2	8	8	10
第四次	3	7	7	11
第五次	6	4	9	9
	23	27	39	51

平均4.6人，已較接近5。

是的。使用隨機抽取法，可消除外在影響。

1-7 如果在習題 1-6 中的 28 位婦女之名單原先即為已知，我們便可利用此一訊息，將兩組之組成份子的婚姻狀態平均分配如下：

- (a) 首先將 28 位婦女分配成 14 對——其中有 5 對為單身婦女（以大寫字母表示者）與 9 對已婚婦女（小寫字母）。
- (b) 就每一配對中之第一位婦女擲一硬幣視其出現正、反面而決定將之指派至處理組或控制組。例如，若硬幣出現「正面」，則將該對之第一位婦女指派至處理組，而另一位則指派至控制組。

依照此一過程指派每位婦女所屬組別，則已婚婦女及單身婦女是否能剛好平均分配於兩組？

解 (a) 單身：A B C D E F G H J K

已婚：a b c d e f g h j k m n p q r s t u

先各將單身、已婚者隨機分成兩兩一組：

單身（先令 $A = 1, B = 2 \cdots$ 利用附表一，從第 7 列第 1 行開始，直找，兩個號碼為一組，重複則跳過）

(8, 9) : (H, J), (1, 4) : (A, D), (2, 6) : (B, F),
(5, 7) : (E, G), (3, 10) : (C, K)

已婚（用同樣方法）：

(a, g), (u, m), (b, c), (p, d), (t, f), (s, k)
(e, h), (n, q), (j, r)

然後擲銅板：

單身：實驗組 H, D, B, E, C

控制組 J, A, F, G, K

已婚：實驗組 a u c p f s h n r

控制組 g m b d t k e q j

- (b) 此法可使單身和已婚的母親得以平均分配至兩組，隨機選擇可避免可能影響結論的因素。

1-8 在 1975 年柏克萊大學自 8300 名男生挑選 3700 名以及自 4300 名女生中挑選 1500 名入學（Bickel 與 O'Connell, 1975）。假設男、女學生在申請入學時，獲得許可之機會均等，則

- (a) 男、女學生申請入學之獲准率有何差異？你認為此一比率是否構成「歧視女性」之證據？

(b)試由下表中找出此一差異的原因。(該表資料雖為臆設，但在簡化計算的原則上，亦掌握了問題之癥結)。

學 院	男生		女生	
	申請人數	獲准入學人數	申請人數	獲准入學人數
文 學 院	2300	700	3200	900
理工學院	6000	3000	1100	600
總 計	8300	3700	4300	1500

試問在文學院及理工學院中，入學許可比率各有何差異？存在「歧視女性」之現象的是那一個學院？

(c)乍看之下，(a)與(b)之答案是否有些矛盾，為什麼？(此即為「Simpson 矛盾」之例：個別部份為真，但若就整體來看則未必為真)。

(d)回答下列各題之「真」、「偽」；若為「偽」的話，請指正其錯誤。

(1)如果不考慮到學院的不同，則男、女學生獲准入學的機會大致相同，然而，有一趨勢顯示，女生大多往較難申請的學院擠，因此整體而言，她們獲准入學的比率偏低。

(2)問題(a)中，入學許可率在性別上的差異是可以理解的，因為存在一外在因素(學院)無法控制，然而，在問題(b)中，我們已將此一因素控制妥當，則問題的真相為何？

(3)如正文所述，採用迴歸分析或可改進某些缺失，例如，除了考慮「學院」的因素外，亦可將其入學考試成績列入考慮(如果，我們一開始不曾假設男、女學生的資質完全相同)。

解 (a) $P_1 = 3700/8300 = 0.4458$

$P_2 = 1500/4300 = 0.3488$

$$\pi_1 - \pi_2 = P_1 - P_2 \pm 1.96 \sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}$$

$$= 0.0969 \pm 0.0178 \quad (\alpha = 0.05)$$

$\Rightarrow$ 有性別上差異(男性比例較高)

(b)文學院：

$P_1 = 700/2300 = 0.304$

8 初等統計學習題及複習題詳解

$$P_2 = 900/3200 = 0.281$$

$$\pi_1 - \pi_2 = 0.023 \pm 0.0244$$

⇒ 在性別上無顯著差異

理工學院：

$$P_1 = 3000/6000 = 0.5$$

$$P_2 = 600/1100 = 0.5455$$

$$\pi_1 - \pi_2 = -0.0455 \pm 0.032$$

⇒ 有性別上差異（女性比例較高）

(c)這是因為男、女人數相差甚多

且女生偏向申請較難的學門

(d)(1)真 (2)真 (3)真

1-9 在本世紀初置有一份未經證實的研究報告指出，人口密度與社會病態（如犯罪行為）存在正相關，例如，人口密度愈高的地區，其犯罪率愈高。但後來經一些較嚴謹的迴歸分析顯示，如果假設其他因素不變，則此種關係並不存在，甚至可能存在很小的負相關。（Choldin，1978或Simon，1981）。

(a)試列舉本例中可能與犯罪率有關且造成人口密度高的其他因素。

(b)下列各題中，試在括弧中選擇正確者：

(1)這個例子乃在說明〔雙重盲目，觀察研究，隨機實驗〕未經證實前，可能會遭到誤解。

(2)特別因〔外在因素，信賴區間，社會學理論〕，將是造成這些（或全部）誤解的主因。

(3)因此，我們可以說，對於人口密度之影響因素會與其他外在因素〔有所偏差，交互影響，多重影響〕。

解 (a)貧窮、種族衝突、社會動亂、青少年。

(b)(1)觀察研究

(2)外在因素

(3)有所偏差

1-10 在一項針對美國城市中心有關犯罪率的觀察研究（Simon，1981）中，顯示自1950年以來，這些市中心的人口密度降低，而犯罪則提

高。準此，我們該如何說明「低的人口密度是產生犯罪的根源」。

解 低的人口密度與犯罪率的增加，只是我們所觀察到的現象，並沒有把其他相關因素考慮進來，也無法作因果關係上的推論。

1-11 試列舉過去（並且一直延續多年）由於未經評估而形成毫無用處（甚至有害）之「處理」。（例如，醫學界上之換血處理）。

解 纏足、孕婦服用沙利邁寶（一種鎮靜劑，後來發現孕婦服用之，會造成胎兒畸形）。

1-12 試列舉目前（且仍將持續下去）由於未經仔細評估（或沒有人瞭解其確實的效果）而為毫無用處（甚至有害）之「處理」。（例如，終身監禁與易科罰金）。

其中，那些例子較易於合理地評估？如何進行？

又，那些例子實在很難予以評估？為什麼？

解 徒刑、核能發電、設立化工廠。

核能發電也許較好評估，優點：節省能源、空間……。缺點：輻射線、海水污染、廢料處理……。

徒刑則較難評估。

1-13 「獲得MBA學位的人，其終生所得將為高中畢業者的兩倍」(Bostwick, 1977)，如果Bostwick教授之研究發現：獲得MBA學位的人，其年所得為高中畢業者之兩倍。試問，該教授之結論是否正確？

解 此說法言過其實，單純的觀察研究不能證明因果關係，也許MBA與具有高企圖心相關，正因為高企圖心才使得所得增加，而非MBA學位本身的關係，用多元迴歸分析資料可減少偏差。

1-14 安全帶是否真能防止意外傷害或死亡？某一實驗中（調查車子擦撞時，虛擬的受傷乘客）證實安全帶之效果。當然，我們也可以實際的情況，觀察車禍發生時，乘客受傷的情形來證實，例如，在一項意外事故的研究中，我們可以調查車禍發生時，繫上安全帶與未繫安全帶之乘客的受傷比率。在此一研究中，為了將其他外在的影響因素減至最少，在以下每一問題中，那一種處理方式較佳？為什麼？

- (a)使用每一次意外事件的資料，或只利用裝設有安全帶之車子發生意外事件的資料？
- (b)利用每一位乘客的資料，或只利用駕駛先生的資料？
- (c)利用造成各種傷害的資料，或將這類傷害依其嚴重性分等級？
- (d)告訴或不告訴醫生，受傷的病患到底是否繫上安全帶？

解 (a)使用每一次意外事件的資料
(b)利用每一位乘客的資料
(c)將這類傷害依其嚴重性分等級
(d)不告訴

1-15 下表是 1970 年代某 n 年內，從美國 n 個州所紀錄之有關繫安全帶與受傷情況的資料（美國交通部，1981）：

在意外事件中乘客使用安全帶與其受傷之情形

安全帶的使用	所有乘客所佔比率	受傷比率	
		輕度傷害 ($2 \leq AIS \leq 3$)	重傷害 ($AIS \geq 3$)
未繫安全帶	85.9	.023	.013
繫腰安全帶	3.9	.011	.009
由腰部至肩膀之安全帶	7.8	.005	.004
情況不明	2.4	--	—

AIS：受傷害的程度

- (a)倘若以上的資料是整個 1970 年代，全國發生意外事件（車禍）傷患之隨機樣本，試建立輕度傷害比率之信賴區間（繫上由腰部至肩膀者），然後再建立未繫安全帶而受輕度傷害比率之信賴區間：（假設 $n = 10,000$ ）。
- (b)在問題(a)中，重新建立兩種情況（繫與未繫安全帶）而受到重傷害比率之信賴區間。
- (c)根據這些資料，你是否可以評估由腰部至肩膀之安全帶的功用？

解 (a)繫上由腰部至肩膀之安全帶

$$\pi = 0.005 \pm 1.96 \sqrt{\frac{(0.005)(1-0.005)}{10,000 \times 7.8\%}}$$

$$= 0.005 \pm 0.001$$

$$\text{未繫安全帶者: } \pi = 0.023 \pm 1.96 \sqrt{\frac{(0.023)(1-0.023)}{10,000 \times 85.9\%}}$$

$$= 0.023 \pm 0.03$$

(b) 繫安全帶者

$$\pi = 0.004 \pm 1.96 \sqrt{\frac{(0.004)(1-0.004)}{780}}$$

$$= 0.004 \pm 0.001$$

$$\text{未繫安全帶者 } \pi = 0.13 \pm 1.96 \sqrt{\frac{(0.013)(1-0.013)}{8590}}$$

$$= 0.013 \pm 0.002$$

(c) 繫安全帶可減少受傷的機會

1-16 (a) 1977年 Arizona 地區 18,000 死亡人口中有 1,440 位死於呼吸系統的疾病。假設這些死者是由 Arizona 一百萬人口中所抽出的隨機樣本，試計算整個 Arizona 地區死於呼吸系統疾病之比例 (π_A) 的 95% 信賴區間。

(b) 已知全美國在 1977 年，1,900,000 死亡人口中有 110,000 死於呼吸系統的疾病，試計算與上題對應之 π 美國的信賴區間。

(c) 以上這些資料顯示：

(1) Arizona 地區死於呼吸系統疾病的比率較高？

(2) Arizona 地區的氣候較差，因而容易罹患呼吸系統的疾病？

解

$$\begin{aligned} \text{(a) } \pi_A &= \frac{1,440}{18,000} \pm 1.96 \sqrt{\frac{\frac{1,440}{18,000} \left(1 - \frac{1,440}{18,000}\right)}{18,000}} \\ &= 0.08 \pm 0.004 \end{aligned}$$

$$\begin{aligned} \text{(b) } \pi_{USA} &= \frac{110,000}{1,900,000} \pm 1.96 \sqrt{\frac{\frac{110}{1,900} \left(1 - \frac{110}{1,900}\right)}{1,900,000}} \\ &= 0.0578 \pm 0.0003 \end{aligned}$$

$$\text{(1) } \frac{0.08}{0.0578} \approx 1.4 \quad (\text{約 } 1.4 \text{ 倍})$$

(2)不一定，可能有其他因素。

1-17 美國的失業人口統計係來自當期的人口調查之資料，這是美國人口普查局每月由 100,000 個成年人口之樣本所做的推估。正如蓋洛普民意調查一樣，它是經由多重階段與其他隨機抽樣所獲得的樣本，其精確性猶如簡單隨機抽樣所獲得之樣本，因此本章 (1-2) 式是一個很接近的推估式。1977 年 12 月份所抽取的樣本（年齡均超過 16 歲）如下：

就業人口	58,770
失業人口	3,730
非勞動人口	37,500
總計	100,000

根據以上的資料，全美國在該月份之失業率為何？並建立其 95 % 之信賴區間。

解

$$\begin{aligned}
 P &= \frac{3,730}{100,000 - 37,500} = 0.05968 \\
 \pi &= 0.05968 \pm 1.96 \sqrt{\frac{0.05968(1 - 0.05968)}{62,500}} \\
 &= 6.0\% \pm 0.2\%
 \end{aligned}$$

第二章 敘述統計

習題

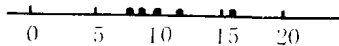
以下有許多習題均是實際資料為依據（如習題 2-1 或 2-2）。為了兼顧實際資料及代數上的簡易處理，我們進行如下之設計：從可獲得的資料（通常是整個母體，或至少是很大的樣本）中，盡可能地建立能夠反映其母體的小樣本。

- 2-1** 在美國的某一所大學中，隨機抽取 5 位女教授為一樣本，其年所得如下：（單位千美元：katz, 1973）

9, 12, 8, 10, 16

試以 X 軸上之點代表每人之年所得，不必分組繪出其分佈圖。

解

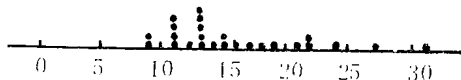


- 2-2** 於上題中，在同一所大學，隨機抽取 25 位男性教授為一樣本，其年所得如下：（單位仍然為千美元）

13 11 19 11 22 22 13 11 17 13
27 14 16 13 24 31 9 12 15 15
21 18 11 9 13

以同樣的刻度，不必分組畫出如同上題之分佈圖，並與上題比較，可知男、女教授年所得之差異。依你的論點看，這些數據是否足以證明全校中男教授之年所得普遍高於女性教授？（本題更精確的答案，可參見第 8 章）。

解



男教授的薪水似乎比女教授高。

- 2-3** 試以 10, 15, 20, 25, 30 將問題 2-2 予以分組，並畫出長條圖。