

萬有文庫

第二集七百種

王雲五主編

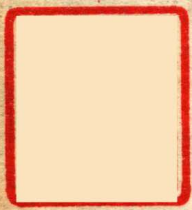
統計學原理

(二)

鮑萊著

李植泉譯

商務印書館發行





統計學原理

(二)

鮑萊著

李植泉譯

陳錦濤校

漢譯世界名著

第七章 圖示法

第一節 總論

平均數與圖式 在初級統計學中，有重要方法兩種，為一般學生或辦理統計事務之官員，所不可不知，且無數學知識亦易瞭解，但竟為一般所誤解，或被對統計無興趣或向不通統計之人士所忽略者，厥為平均法與圖示法。此兩種方法所以相提並論者，乃以平均數及圖式之用途，有幾不可分之連繫在也。設遇有極大極繁之數字羣類當前，雖用表列法已為明晰之表示，但吾人對於許多數字之整個情形，仍不能完全領略。任何一系列數字——各城市之人口，各級年齡上之死亡率，許多工人之工資，若干年數之進口貨值——數列愈長，愈難令人索解。十個數目組成之序列，吾人或可一望而知，二十個數目之序列，雖稍費力，亦不難明瞭；但如一系列印就之一百年數字，則欲其予吾人以何等之印象也殊難。一樹易見，成林之樹木則不易辨清矣。關於平均數用途上所有問題之試金石，即視所選用之平均數，是否能給予全部羣類以最佳之總結，使此一總結數字令人可以一目瞭然。在平均數一名詞意義大為擴張時，即可知同時選用三，四乃至十個適當之數字，

悉能充分表現任一羣類之主要形態。平均數如此，圖示法亦然，圖示法之主要用途，乃在表現大羣類之數字，以便瞭解全部之情形，至一切圖式之是否完善，其試金石乃視所繪之圖，能否托出一組數字之最佳觀感，令人一望即可領會。惟圖式尚有一用途，為平均數所不及者，即欲表現時間數列，唯圖式實優為之。然就實質言之，圖式又不如平均數之重要，蓋平均數雖由若干數字推出，但離開數字亦能獨立存在，而代表所測量之數量之真正形態。唯用平均數，乃可以此一羣類，與其他之羣類相比較，而圖式則不然，蓋圖式乃立於輔助之地位，非主要之元素，即完全取消之，亦無不可，因用之乃在助目力之所不及，且為閱者節省時間計也。

○~~~~~○
 {圖式法及平均數} 為使本章與前一章關係銜接起見，茲請同時用兩種方法，表示同一之數字羣類，即以多數工人之工資為例以討論之，資料如下：

工人之工資及人數

工 資	人 數
自十五先令至十六先令	200
自十六先令至十七先令	400
自十七先令至十八先令	100
自十八先令至十九先令	100
自十九先令至二十先令	200

} 1000

自二十先令至二十一先令	200	} 2,200
自二十一先令至二十二先令	300	
自二十二先令至二十三先令	300	
自二十三先令至二十四先令	500	
自二十四先令至二十五先令	900	
自二十五先令至二十六先令	1,200	} 3,500
自二十六先令至二十七先令	800	
自二十七先令至二十八先令	700	
自二十八先令至二十九先令	500	
自二十九先令至三十先令	300	
自三十先令至三十一先令	300	} 2,100
自三十一先令至三十二先令	400	
自三十二先令至三十三先令	400	
自三十三先令至三十四先令	500	
自三十四先令至三十五先令	500	
自三十五先令至三十六先令	600	} 1,200
自三十六先令至三十七先令	400	
自三十七先令至三十八先令	100	
自三十八先令至三十九先令	80	
自三十九先令至四十先令	20	

如用平均法，上列羣類可以下列數字代之：

全部之平均數 二十七先令六便士

最低1000人之平均數一十七先令

最高1000人之平均數三十六先令六便士

中等4000人之平均數二十七先令

或

中位數 二十六先令九便士

四分位數 二十四先令二便士

三十二先令

十分位數 二十先令

二十三先令六便士

二十四先令九便士

二十五先令八便士

二十六先令九便士

二十八先令二便士

三十一先令

三十三先令四便士

三十五先令四便士

衆數 二十五先令三便士

次要衆數 十六先令六便士

三十六先令

或

工資

人數對全部之百分數

自十五先令至二十先令

10

自二十先令至二十五先令	22
自二十五先令至三十先令	35
自三十先令至三十五先令	21
自三十五先令至四十先令	12

{簡單圖式之構造法} 此一羣類，以圖式表列之，如第二圖。

此為用圖表示兩變數相互關係之一例。與此相類之圖式，可用以表示各級年齡中之結婚率或死亡率，表示身長各異之人數，在各種價格上之需要，或用以表示其他任何同質數量之羣類。即以此同一之構圖法，表示若干年中任何數值之變動，亦無不可。構圖之法，須在紙之底邊，畫一平行線，在此線上分成相等之距離，以代表依次增加之若干數量，諸如年齡，所得，身長，價格，時間以及其他等等。此線名曰橫軸 (axis of abscissa)，線上某一點至零點之距離，謂之該點之橫坐標 (abscissa)。在紙之側邊，經過零點再畫一線，與橫軸成爲直角，則此垂直之線，即名縱軸 (axis of ordinate)。在此縱軸上，亦分成等距，依次代表含有橫軸上所代表數量之人數或件數。又於橫軸上分別在各點各畫一垂直之線，標明在該一點上有若干人數或件數，此線即爲縱坐標 (ordinate)。請閱下列一圖，橫坐標所表示者，爲工資數，縱坐標代表賺各級工資之人數。橫坐標縱坐標既已判明，然後即在各縱坐標之上端，各連一直線，於是圖式乃告完成。實際上，如用方格

紙，則可不用畫縱坐標，而縱坐標之頂端，亦可標出。

工資圖之解釋。 由此一圖，令人一望即知，賃銀勞動者依所賺工資分配之情形。工資在十五至十六先令者人數甚少，在十六至十七先令者則漸多，但在十七至十九先令者人數又大為減少。自十九先令起，人數依次連續上升；至二十四先令與二十七先令之間，人數乃登峯而造極，其中尤以在二十五先令至二十六先令者為最多。但自是而後，則每況愈下，直至三十先令一組而後已。此三十先令以前一組人數雖過少，而與十七至十九先令之間者相較，尚非過低。降至三十先令之後，復又以有規則之進展，漸次上升，至三十六先令後，乃又急轉直下，而達於三十九先令一組之最低數。由此觀之，在二十五先令左右集合之大部，乃為全類之主體，而三十六先令之一組則屬其次，但在十六先令之處，則為極小而幾別成一部。凡此種種，示與吾人，在三十至四十先令間，有高度技能之工人甚多，至有平常技能之工人大眾，則其工資乃在二十先令至三十先令之間，但有極少之粗笨無技能工人，工資均在十六先令左右。用一圖式形態畢現，如用表列豈可得哉！

於此吾人須加注意者，表上在十五先令至十六先令間之人數，正由在十五先令六便士（該組之中點）之縱坐標表示之。如表列所根據之原來數字，能取至最近之一便士為止，則縱坐標應盡在十五先令五又二分之一便士處。各組之中點位置，畫時須準

確安置，此一要事也。

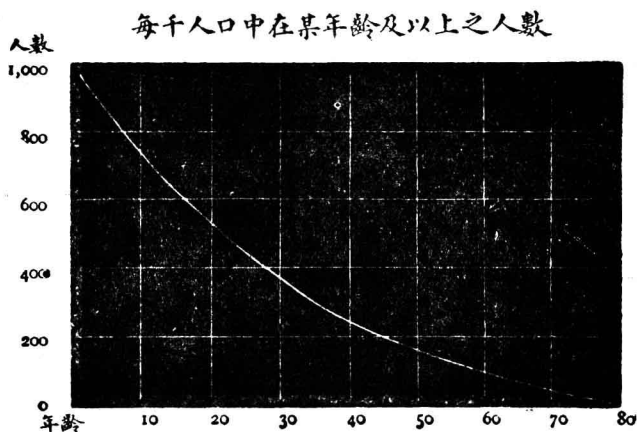
連續性 連結各縱坐標頂端之線，用途有二。第一爲便於判斷各相對高度 (relative height)；第二可表示連續性 (continuity)。第一點姑不具論，茲就第二點作圖舉例解明於下：

每千人口中之年齡分配

(根據一八九一年之人口普查)

年齡	人數	年齡	人數
自零歲以上	1000	自十七歲以上	607
自一歲以上	973	自十八歲以上	587
自二歲以上	949	自十九歲以上	567
自三歲以上	925	自二十歲以上	547
自四歲以上	901	自二十一歲以上	528
自五歲以上	877	自二十二歲以上	510
自六歲以上	854	自二十三歲以上	491
自七歲以上	830	自二十四歲以上	474
自八歲以上	807	自二十五歲以上	456
自九歲以上	783	自二十六歲以上	439
自十歲以上	760	自二十七歲以上	423
自十一歲以上	738	自二十八歲以上	407
自十二歲以上	715	自二十九歲以上	391
自十三歲以上	693	自三十歲以上	376
自十四歲以上	671	自三十一歲以上	361
自十五歲以上	649	自三十二歲以上	346
自十六歲以上	628	自三十三歲以上	332

年 齡	人 數	年 齡	人 數
自三十四歲以上	318	自五十八歲以上	35
自三十五歲以上	305	自五十九歲以上	79
自三十六歲以上	292	自六十歲以上	73
自三十七歲以上	280	自六十一歲以上	67
自三十八歲以上	268	自六十二歲以上	62
自三十九歲以上	256	自六十三歲以上	57
自四十歲以上	244	自六十四歲以上	52
自四十一歲以上	233	自六十五歲以上	47
自四十二歲以上	222	自六十六歲以上	43
自四十三歲以上	211	自六十七歲以上	38
自四十四歲以上	201	自六十八歲以上	34
自四十五歲以上	191	自六十九歲以上	31
自四十六歲以上	180	自七十歲以上	27
自四十七歲以上	171	自七十一歲以上	24
自四十八歲以上	161	自七十二歲以上	21
自四十九歲以上	152	自七十三歲以上	18
自五十歲以上	143	自七十四歲以上	15
自五十一歲以上	135	自七十五歲以上	13
自五十二歲以上	127	自七十六歲以上	11
自五十三歲以上	119	自七十七歲以上	9
自五十四歲以上	112	自七十八歲以上	8
自五十五歲以上	104	自七十九歲以上	6
自五十六歲以上	98	自八十歲以上	5
自五十七歲以上	91		



(第 三 圖)

上圖，橫軸代表年齡，縱軸代表一八九一年夏末英格蘭及威爾士人口每千人中尚存在及在各級年齡以上之人數估計數。縱坐標乃由代表每歲年齡之中點之橫軸量起；但壽命之長短，不能照幾年幾月幾日計算也。此圖之本意，即為表明每一整歲尚存人口之比例，為達此目的計，縱坐標之頂點所連成之線，必須不致出現破斷及折角，而應有絕對之連續性也。

實際上，不滿一整歲之較小年齡組距，無標出其縱坐標之必要，恐令人一見不能領會其詳細情節也。然上圖所畫之線，亦確與人口數為無限大而年齡分組無限小者所呈現之線之形式相同。

就此所論各點，請考察前繪之工資統計圖（第二圖）。某一年之平均工資收入，計數時未必每一先令，甚至每一便士，均無

錯誤，此乃觀察件數不足之故，苟如件數充足，吾人自可將工資分爲整齊之順序，而以微至一法尋(farthing)爲準，則表示工資之線，必不致有尖銳之折角，而成爲連續之曲線。圖式所給予吾人之印象，即爲連續性之存在。又吾人於下列第六圖中，可得一印象，出口貨值之線，亦確爲日接一日連續不斷也。

面積 經過一顯然之階段，吾人頗可假定，由橫軸相接兩區分點畫成之兩條垂直線，與由縱軸相接兩點畫成之兩條水平線，四條線交叉所成之中間方格面積之單位，乃代表一個工人；由此假定，可知上以曲線爲界，下以基線爲止，以左右兩條經過代表任何兩種工資額之兩點所成之垂直線爲範圍，所畫成之面積，即爲工資介乎底線上所代表之該兩種數額之工人總數。

故在第二圖中，經過中位數， M ，四分位數， Q_1, Q_3 ，十分位數， $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9$ ，等點所繪成之線，即將面積 $ABm_1m_2m_3CD$ ，分別分爲二，四，及十個等面積。該圖之重心，乃在於經過平均工資， V 所畫成之垂直線上；至經過最高位置 m_1, m_2, m_3 三點之縱坐標，在基線上之基點，即爲三個衆數之所在。

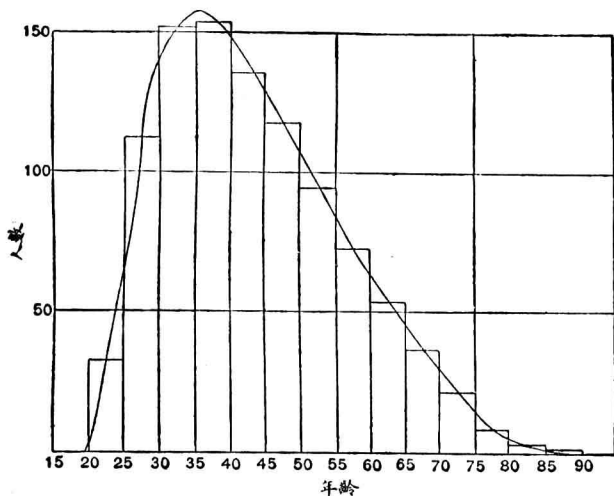
分組資料 在資料之分組甚寬時，吾人以用下列方柱圖(block diagram)之法爲宜。

一九一一年英格蘭及威爾士已婚男子之年齡分配

甲 方柱圖——每千人中每五歲一組之人數

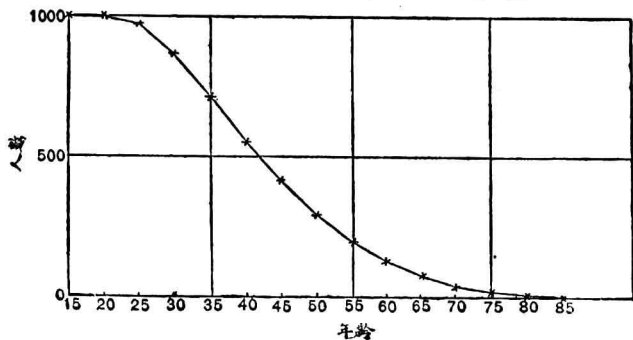
第四圖 一九一一年英格蘭及威爾士已婚男子之年齡分配

甲、方柱圖—每千人中每五歲一組之人數



(第四圖甲)

乙、累積圖—每一千人中各級年齡以上之累積人數




(第四圖乙)

上列之甲，乙，二圖，乃為前表第十八表已婚男子人數依年齡分配之例釋。在第十八表中，吾人除只知其年在二十歲而不滿二十五歲之比例外，並無別項消息可得。但今用圖式，以代表五歲之組距為底，以與該組記錄上人數成比例之數為高，乃成一長方形，所有前表未能宣示之事實，茲均已精確之表示。根據前第二圖之指示，各組人數乃均在各該組之中心。如就年齡之圖式而論，各歲人數理應連續不斷，於是如完全表示之，即為連續曲線，吾人已熟知之矣，然則以五歲組距為底之面積，必與各該長方形之面積相等。此種曲線雖為用隨手畫法，在圖上繪成，但若全線之位置，並無動搖不定之處所，則此曲線必足為事實之代表。

資料亦可用圖乙代表之，在圖乙中，十字號即為表上記錄之資料。將此若干十字號各連成直線，若事實現象果為連續的，則結果所得之線，與曲線乃無二致，即在本圖之曲線，亦殊難與各點所連成之直線辨別也。

必需之確度 繪圖之詳細技術，標尺尺度之位置，使圖式明晰之辦法，等項可於本章所舉各圖示明之。至於數字究以萬為單位，以千為單位，或以一個為單位，其確實之程度，乃純以目力能以領會為決定之標準。在本章所舉之圖例，所表示之數字，設在一千個中，有一易其位置，即判然可見，此乃平常之限度也。若進

而欲求更高之確度，誠亦非事實所需要，蓋圖式不能離表列而獨立，其功能亦僅為補目力之所不及，特別顯示數列之重要形態而已。

 在未討論選擇代表數字之尺度前，圖式必如何乃可予吾人以明顯之印象，應有加以研究之必要。夫吾人之肉眼，所能判斷者有三：一，距離；二，比率；三，角度是也。茲以第六圖之虛線為例，討論距離，比率及角度三端。

(一) 肉眼對於距離之長短，可予以頗為安全之判斷；兩點對於基線之距離，何者較遠，何者為近，一見即知，幾無可疑之點；若用方格紙，則雖有千分之一之差，亦能明察無遺。且眼目對於差額之判斷，最為神速；例如圖中一八八三年出口貨值超過一八八五年出口貨值之數，較之一八九〇年出口超過一八八三年者為尤巨，一望可知也。

(二) 一八八九年之出口貨值，幾為一八六二年之兩倍，或一八七八年之出口貨值，約當一八九〇年之四分之三；均不難一望而知。但只恃目力，其測量所得之確度，並不甚高；一八七三年出口貨值與一八七一年出口貨值之比(1.095:1)，較一八八二年出口貨值與一八八〇年出口貨值之比(1.073:1)為大，實難用目力覺察之；然此並非目力不能判斷之證明，反之，該圖所給予吾人之印象，一部乃由此種性質無意識之計算而完成也。至如欲

得確實之觀察，須用本章第五節所述之方法。於此，吾人須注意，爲表現此等觀察，必須插進基線；且因目力之計算，本爲無意識者，若以一圖表示若干年之運動形勢，而無基線，則其將給予吾人以錯誤之印象也無疑。

（三）設吾人質問：一八八六至一八八七年增加之量大耶？抑一八八七至一八八八年之增額大耶？則欲得迅速之答覆，與其查視二年度之差額，無寧觀察二者上升之角度。後一年度之變動，上升之線，較前一年度上升之線爲峻峭（與水平線成一較大之角）；故後一年之變動，必較前一年度者爲大，實際上，後一年度增加有一千二百六十萬鎊之多，而前一年度只增加九百二十萬鎊也。此種目力觀察最有效之練習，莫過於判斷增加率發生變動之年度；例如，一八六二至一八六三年，出口貨值既有增進，一八六三至一八六四年，增加漸緩，一八六四至一八六五年漲勢仍輕，在一八六五至一八六六年，則驟然上騰；自一八六六至一八六七年，漸漸下降後，復又以加速度接續上升，直至一八七一年而後已，其他年度依此類推，姑不多述。一八七二年至一八七六年之線，與基線成凹面（concave），乃爲急轉直下之形勢；一八七九至一八八二年之凹形，又緩緩上升。總之，如此所示之增加，乃絕對或實際之高漲，並非與各該年度之最初數量成比率而計算也。

尺度之選擇 尺度爲定數字位置所必需，究應如何選擇之，殊難規定準則。吾人所欲研究者，只縱橫尺度之相對比率一問題。全圖不可過大，以一望可窺及其全豹爲宜；如圖式複雜，所述者爲極多年代，及極多之變動數字，則纖細之確度，乃不得不付之犧牲。設橫標尺尺度，業已定妥，選擇縱標尺尺度之法，必須使一代表增加率最大之線段，能表現對縱軸十分傾斜之形勢爲度，而欲其如此，唯有將縱標尺尺度定爲極小之一度。但若欲使所有數字重要之異動，顯然可見，則縱標尺尺度不必加大。唯能適合此兩種條件之尺度，乃可謂之適得其當。下列之第五圖，例示由故意操縱尺度之大小，及取消基線，所能給予吾人之錯誤印象。

正當底線之必要 此組粗略之圖式，均表示同一之數字：一八六〇至一八九一年英美兩國工資估計數。圖甲之各線，尙爲正當。圖乙，英國工資數之尺度，基線並非自零點起算，而美國數字亦大爲縮減；其結果乃成英國工資變動甚大，而美國數字略有增加之情形。圖丙，丁，戊三圖，尺度大加改竄，基線亦行更易，故圖丙之美國工資，有超出英國數字而上升之勢，而圖戊情形又恰形相反，同時圖丁，兩國之工資，乃成等速運動之情狀。試就以下各圖，加以分析，足見只憑目力決定基線，基線必難正當無誤，若取消基線，又必不能據以估計變動之重要性。又除少數之例外容當提