

統計學

金國寶編著

臺灣商務印書館印行

特

號

•王雲五主編•

人人文庫

統計學

金國寶編著

臺灣商務印書館發行

復刊人人文庫序

人人文庫自民國五十五年始刊，迄六十二年終刊成者計達一千五百餘種。中分單號雙號及特號三種。單號每冊八元，雙號十二元，特號二十元。其種數之多，定價之廉，冠於全國。及六十二年秋後，紙張價格奇漲，且不易得，其他工料莫不稱是。人人文庫原以廉價為主，隨成本而增價，殊違本旨，不得已於六十三年元月始暫停新書之印行，即原已出版各書亦以售罄爲止，暫不重版。今歲三四月以來紙價工價雖平均較前增長百分之一百五十，然已漸趨穩定，籌謀再四，決從五月起，仍予復刊，每月新刊暫定爲十種，其原出各書，銷數較廣者，仍予重版，以應讀者需求。書價姑定爲單號每冊十二元，雙號十八元，特號三十元，所增雖僅百分之五十，而以視工料之增長百分之一百五十者，仍稍虧損在所不惜。

復刊以後，選材益加審慎，範圍亦日廣，除與英國之人人文庫比擬，且後來居上。關於新知識之介紹仍略仿英國家庭大學叢書。又復刊新書之編著悉與原刊蟬聯，設印刷工料不再增長，則由原刊之千五百餘種，不難與時並進，遞增至數千種，乃至萬種，使青年學子得以廉價盡讀有用之書，此則所殷望也。

又除單號雙號每種仍維持一冊外，特號因多載名著，爲存其真，必要時得分訂爲二冊以上，如十九世紀歐洲思想史卽其一例也。

中華民國六十三年五月一日王雲五識

目次

第一章	緒論	1
第一節	統計學之定義	1
第二節	統計學之應用	2
第三節	統計之誤用	3
第四節	統計之法則	6
第五節	統計方法之程序	7
第二章	統計表	9
第一節	統計表之功用	9
第二節	統計事項之特性及其相互之關係	9
第三節	統計事項之分類	10
第四節	總表及摘要表	11
第五節	統計表之形式及製表規律	12
第六節	統計數列	15
第七節	頻數表	17
第三章	統計圖	25
第一節	統計圖之功用及製圖之原則	25
第二節	統計圖之分類	25
第三節	條形圖	26
第四節	統計地圖面積圖及體積圖	34

第五節	線圖.....	37
第六節	作圖規則.....	45
第四章	平均數	51
第一節	平均數之意義與種類.....	51
第二節	算術平均數.....	52
第三節	中位數.....	63
第四節	衆數.....	71
第五節	幾何平均數.....	75
第六節	倒數平均數.....	78
第五章	離中趨勢與偏態	87
第一節	離中趨勢之意義及其測定之方法.....	87
第二節	全距.....	88
第三節	四分位差.....	88
第四節	平均差.....	90
第五節	標準差.....	95
第六節	各種離中差之關係.....	102
第七節	偏態之意義及其形式.....	104
第八節	測定偏態之方法.....	104
第六章	指數	113
第一節	指數之意義與種類.....	113
第二節	物價指數編製之方法.....	113
第三節	指數公式之測驗.....	126
第七章	吾國重要指數之編製	133
第一節	物價指數.....	133

第二節	生活費指數.....	137
第三節	外匯指數.....	146
第四節	證券指數.....	148
第五節	國外貿易指數.....	152
第八章	繫聯	155
第一節	繫聯之意義.....	155
第二節	繫聯係數之計算.....	156
第九章	時間數列	169
第一節	經濟現象變動之原因.....	169
第二節	長期趨勢.....	169
第三節	季節變動.....	170
第四節	循環變動.....	174
第五節	商情預測.....	176
第十章	機率與差誤正態曲線	180
第一節	機率.....	180
第二節	差誤正態曲線.....	190
附錄甲		201
美華對照統計名詞.....		201
附人名地名索引.....		214
附錄乙		217
統計符號.....		217

附錄丙	219
-----------	-----

本書重要參考書	219
---------------	-----

附錄丁	221
-----------	-----

計算應用表	221
-------------	-----

統計學教科書

第一章 緒論

第一節 統計學之定義

統計學者用計數或估量以數字表示社會或自然現象之動態或靜態並分析其數字間關係之學也。此定義須稍加以解釋。

統計學計量而不較質。欲比較人之貧富或智愚，在統計學內必須先有可以表示此貧富或智愚之數量方可以言比較，故數字不能與統計分離。統計學上大半數字均由計數而來，故雷翁衰氏以計數之學作為統計學之定義。此定義雖覺太狹，但計數為統計學之主要職務要無可疑。惟統計學上之數字未必均由計數而來，有時不得不用估量方法以求其近似之數值，故定義中計數與估量並列。

古代統計學研究之對象為國家，故有以研究國家之學為統計學之定義者。其後研究之範圍漸次推廣，研究之對象亦漸由國家而推及於社會與自然現象；此種現象或同時同地，或同時異地，或同地異時，故社會與自然現象之動態與靜態均在統計學研究範圍之內。

統計學亦有作為研究平均數之科學者。此定義亦覺太狹。統計學不特用數字表示社會與自然現象之動態與靜態，且用種種分析方法以推求其數字間之關係；此種關係不僅是平均數一種。故謂統計學之任務在根據大量觀察而闡明其數字間之平均關係則可，若謂統計學為研究平均數之學，則不免令人誤解矣。

第二節 統計學之應用

統計學之應用甚廣，不勝枚舉，茲擇其最重要者分述如下：

(一)統計與社會政策 近世各國無不倡言社會政策，社會立法；然欲救濟社會之疾病，必須先明瞭社會疾病之原因，然後對症下藥，方可有改良社會之望。欲明原因則非取證於統計不可。例如根據工業上之失事統計，於是有強迫保險之實行；比較男女童工之工資，知女工童工有特別保護之必要，於是有最低工資法之制定，即其例也。

(二)統計與公共衛生 人口之疾病死亡統計對於公共衛生尤有密切之關係。衛生當局之唯一參考即在統計；凡以後施政之方針及以前設施之成績無不取決於此。當疾病發生之際即可由統計之報告而設法防止其蔓延；平日亦可注意社會之弱點而徐圖補救之辦法。且若將此等統計及其效用公告人民，尤不難得社會之合作與經濟之贊助，誠推廣公共衛生事業之第一急務也。

(三)統計與商業 現代商業範圍擴大，故其問題亦日趨複雜；內部如浪費之減少，工人之效能，分公司之營業，售貨員之比較，等等；外部如供給需要之狀況，市場之變遷，商業之盛衰，季節之影響，等等；皆與商業之成敗，有莫大之關係。故現代歐美各大公司皆特設統計部以專司其各種調查之職責。

(四)統計與財政 財政以收支適合為原則；支出雖較能預定，然收入卻頗難預言。例如所得稅之多寡，須視人民所得之數額而定；關稅之收入，須視外國輸入品之種類與數額而定；然歐美各國之財政專家每能根據歷年之統計而預測未來之收入，雖亦有時與實收數目相差甚遠，然適合者其常，而相差極大者僅例外耳。

第三節 統計之誤用

統計之用固極神妙，然用之失當，其流弊所屆亦有不可勝言者。吾人格物務須平心靜氣屏除成見，取懷疑之態度，戒獨斷之行爲，如是細心分析方能得事物之真相；否則，毫釐千里未有不陷於絕大之謬誤者。美國統計學家卻獨克氏對於統計學之誤用論列甚詳，據其所論可分謬誤爲四種：一曰，不同事物比較之謬誤；二曰，百分比之謬誤；三曰，原因脫漏之謬誤；四曰，偏見之謬誤；而統計自身之謬誤猶不與焉。卻氏對於每種謬誤各附以若干例證，茲節取其一二如下：

（一）不同事物比較之謬誤 統計之妙用端在比較，單獨一個數字毫無意義可言，必有兩處地方或兩個時期相互比較，意義始明；然事物之性質不相同者則無比較之可能。如一八九九年美國陸軍部長論菲列濱之美國兵士死亡率事，即犯此病。當時外界對於多數兵士之死亡頗有責言，該部長乃出而置辯；大意謂兵士之死亡率不過萬分之一七二，與華盛頓波士頓一般人口之死亡率相差無幾，故兵士之死亡率不得謂爲過高云云。其實軍隊與一般人口，性質完全不同，安能相提並論？一般人口之中，老少齊全，而極老極少者之死亡率尤高出尋常，斷不可與中年人相比較；而兵士則既悉強壯之青年，且均經過體格檢查者，故此種比較實自欺欺人耳。

又如就各國煤礦工人每千人每年所遇之失事數目而比較之，其結果亦不確當；蓋各國礦工工作之日數不同，歐洲各國較多，美國較少。若以每年所遇之失事數相比較，則美國必佔便宜，故最善之辦法須將各國之失事次數均以作工三百日爲標準而修整之，方有比較之可能。

(二)百分比之謬誤 百分比之使用亦須格外注意，偶一不慎即可令人發生謬誤之感想。例如美國約翰哈金斯大學初收女生之時即發生一有趣之新聞，謂該校女生百分之三十三又三分之一均與本校教員成眷屬云。不知底細者必以為該校教員均風流人物，然細加考察則與教員結婚之女生僅一人而已。蓋當時女生共祇三人，一人即為全體百分之三十三又三分之一。故數目甚小之時，不宜用百分比。如欲用百分比，亦必須將實在數同時並列，方不致令人發生謬誤之印象也。

又如甲城人口十萬而外國移民居百分之二十，乙城人口五十萬而移民居百分之三十，丙城人口百萬而移民居百分之四十，今若將此三城合而計之，試問其移民成分幾何？對此問題往往即將此三城之百分比相加而以三除之為答，則其答數為百分之三十：

$$\frac{20 + 30 + 40}{3} \% = 30 \% ;$$

然應得之百分比當為 35.625 而非 30 也；因準確之計算法當以三城人口之總數除其移民之總數。甲城之移民有二萬，乙城之移民有十五萬，丙城之移民有四十萬，故三城移民之總數共有五十七萬，而其人口之總數則有一百六十萬，列成算式即得：

$$\frac{570,000}{1,600,000} = 35.625 \%$$

(三)原因脫漏之謬誤 有時事實之原因甚多，若獨取其一盡置其他於不顧，亦常發生不確之結論。如美國某大學調查學生吸煙程度以斷定吸煙為學業不及格之原因，即其一例。茲將學生分為吸煙多者，吸煙不多者，完全不吸者三類，而其中不及格之人數如下：

	調查學生數	全年平均分數	不及格之百分比
完全不吸者	111	85.2	3.2
吸煙不多者	35	73.3	14.1
吸煙極多者	18	59.7	24.1

根據此項調查結果遂認吸煙爲不及格之原因，則未免失當；蓋學生之好吸煙者其人往往視他種活動較重於學業，同時或爲體育家，或爲極貪舒適之人，其所以不及格者由於其不重學業所致，而吸煙不過其不重學業之一種間接表示。科學家之探討必須盡窺事物之全豹，不當以部分的理由作全體之解釋也。

(四)偏見之謬誤 統計學家當有超然中正之態，切不可先有偏見存乎胸中乃覓統計以實其說；自欺欺人，莫此爲甚。例如前年美國嘗有反對種痘同盟會之職員投函於紐約晚報（一九一四年五月四日），引用英國之統計如下：

天花死亡總數 (1905—1910 年)	199
種痘死亡總數 (1905—1910 年)	99
五歲以下天花死亡總數 (1905—1910 年)	26
五歲以下種痘死亡總數 (1905—1910 年)	98

大意謂就全體而論，種痘死者歲佔天花死者之半，就五歲以下而論，種痘死者尚多出天花死者之上幾有四倍之多，故強迫種痘之舉殊可不必云云。所引統計固屬確實，但其議論殊與事實相反。天花死亡之減少即由於種痘。六年之中，因種痘致死者祇有九十九人。假使不實行強迫種痘，則當時以天花死者必非少數，以之與九十九人相比，孰多孰少，不言可知矣。

然而以上種種猶非就統計本身言也。統計本身亦難免謬誤，或由於調查之疏忽，或由於計算之錯誤，故於他人所披露之統計吾人須詳加分析，不可輕信。差以毫釐，謬以千里，不可不

慎之又慎也。

第四節 統計之法則

抽樣爲近世所發見最有價值之調查方法。所謂抽樣即自一大羣極複雜之事項中抽取一小部分作爲調查之標準，由此所得之結果即可用以代表全部。例欲調查上海工人所得之平均工資，吾人不必遍查全部工人所得之工資再求其平均數，吾人祇須抽查其中可以代表全體的一小部工人所得之工資再求其平均數。由是而得之平均工資雖未必與全體工人所得之平均工資完全一致；然相差甚微，實際上可以略而不計，故即以之作爲全體工人所得平均工資之代表亦無不可。又設有雞蛋十萬枚而欲求其平均重量，吾人不必將此十萬枚雞蛋一一秤其重量再求其平均數。吾人祇須任取（當然不能故意選擇最大或最小之雞蛋）其中一千枚雞蛋秤其重量而求其平均數。由是而得之平均重量雖未必與十萬雞蛋之平均重量完全一致；然相差無幾，實際上已可用爲全部雞蛋平均重量之代表。此種調查方法係根據統計常態之法則。所謂統計常態，即謂由一大羣中任意選擇之一小部平均差不多可以保持全部之特性。

由一大羣中任意抽出之一部既能代表全體，則由此一大羣中抽出之其他一部自當與第一部相似。若第一部中有幾項具有異常之特性，則在第二部中吾人亦可預期發見具此異常特性之幾項，其項數亦與前無甚差異；此即所謂小數永存之法則，蓋由統計常態之法則脫胎而來也。統計學家蒲蘭謂各種職業之專家自專醫特種難症之耳科醫生以至販賣古董之商人，靡不賴此「小數永存之法則」而生。

由統計常態之法則脫胎而來，尚有大量惰性之法則。所謂大

量惰性，乃謂在外界原因不變之情況下若觀察之範圍擴大甚廣則每年之統計常得相似之數量。例如火災之損失就一城而言歷年之損失或相差甚多；然就全國或全世界而言，若房屋之建築或防火之設備未有改進，則每年火災之損失常能保持一定之數量；蓋各處每年火災損失數量之變動其方向不同，有較去年增加者，亦有較去年減少者，甲乙等地增加之量適與丙丁等地減少之量約略相抵，故其結果變動甚微。

第五節 統計方法之程序

統計方法之程序，可分為四大步驟：

(一) 搜集資料

(二) 整理資料

(三) 發表資料

(四) 分析資料

請舉例以明之。今設欲清查某地之人口，自當首先確定其清查之範圍，調查表之問題及格式，施行方法，以及其舉行日期，等等，此皆屬於其初步「搜集資料」之工作。迨資料既已齊集，則當進而點明其人口之總數，男女各若干？已婚與未婚者各若干？識字與不識字者又各若干？其年齡之分配如何？其職業之分配又如何？類此之工作皆屬於其第二步「整理資料」之範圍。今既得其統計各事項之數目，即得酌量情形製成圖表以公布之，此即其第三步「發表資料」之工作。大半統計機關之工作至此為止。完成此三項工作亦可謂為已盡「計數」之能事。然統計學之效用卻不但示吾人以各種事項之確切數目，尤當由其所示之數目間發現一定之規律。例如由各國人口年齡統計之比較而發現其分配曲線大致有一定之形式，再若由其已婚者年齡之研究而得計算其夫婦

間年齡之繫聯係數，諸如此類之工作概屬於其第四步「分析資料」之範圍。規模較大之統計機關亦甚努力於此種工作。本書第二、第三兩章先述圖表之繪製，第四章以下略述統計資料之分析，至於統計資料之搜集與整理，在初學者每覺過於枯燥，故從略。

問題及習題

1. 試述統計學之意義。
2. 解釋下列各名詞：
 - a. 計數。
 - b. 估量。
 - c. 抽樣。
3. 統計學之應用何在？
4. 統計之誤用最重要者有幾種？試詳述之。
5. 解釋下列各法則：
 - a. 統計常態之法則。
 - b. 小數永存之法則。
 - c. 大量惰性之法則。
6. 試述統計方法之程序。

第二章 統計表

第一節 統計表之功用

統計不能與數字分離，故其結果恆有無數複雜之數字。此無數複雜之數字即所以表示統計事項之動態或靜態。若用文字一一爲之披露，則長篇累牘，讀者需時既多，而讀後恐仍不知其所云；反之，若將此無數複雜之數字擇要列之於表，則統計之結果便可一目了然，較之用文字敘述者不可同日而語。茲將統計表之功用擇其重要者分述如下：

- (一) 統計資料之排列，有明顯且合於邏輯之系統。
- (二) 易得明切之概念。
- (三) 易於記憶。
- (四) 便於比較。
- (五) 易於檢查錯誤及遺漏。
- (六) 免去文字上重複解釋之煩。
- (七) 便於總計平均及其他較深之計算。

第二節 統計事項之特性及其相互之關係

將雜亂無章之統計資料依一定之系統排列成表，必須先有預定之目的，然後能有整齊之秩序；然欲預定目的，必須先能確定統計事項之特性。所謂特性，即其個別之性質是也。試就田地而言，土質之肥瘠，面積之大小，產量之多少，市價之高低，以及