

大學叢書

統計學

上冊

陳超塵編著

臺灣商務印書館發行

C 8
806
1

S 017178

大學叢書
統計學
上冊

陳超塵編著



S9004284

臺灣商務印書館發行

中華民國七十三年十二月三十三版

大學叢書

統計學

上册

上册

上册

精裝本基本定價四元二角正
平裝本基本定價三元二角正

版權所有
翻印必究

編著者

陳

超

塵

發行人

朱

建

民

印刷及
發行所

臺灣商務印書館

股份有限公司

臺北市重慶南路一段三十七號

登記證：局版臺業字第〇八三六號

序

統計學是一種科學方法，其主要功能為搜集、整理與分析統計資料並推論統計分析結果，其目的在尋求宇宙間之通理通則並建立科學定律。由於統計方法具有此種特殊功能，故其應用範圍極為廣濶，其分析對象幾包括所有經驗科學。所謂經驗科學(Empirical science)即是需要經驗事實證實假設而建立科學定律的科學，在科學領域中，除數學、哲學等少數非經驗科學外，其他各種社會及自然科學幾全為經驗科學。故在社會科學中如經濟學、社會學、教育學等，自然科學中如生物學、心理學以及其他工程及農業應用科學等，無一不以統計方法為其主要研究工具。所謂『工欲善其事，必先利其器』，欲開拓科學領域，進行研究工作，統計方法殆成為必要而有效的研究工具。二十世紀以還，統計學無論在理論與應用方面均有長足之進步，故預料統計方法在未來科學研究上將佔有更重要之地位。

本書共分上下兩冊，上冊專論記述統計學(Descriptive statistics)，下冊專論推測統計學(Statistical inference)。所謂記述統計學即以大量觀察為基礎之各種統計方法，屬於這方面者計有統計資料之搜集、整理及分析等方法，在分析方法中包括平均數、離差、指數、時間數列分析、簡單相關及複相關等，以上各種方法本書上冊均有簡要之介紹。所謂推測統計學即以機率論為基礎根據樣本(Sample)結果以推測母羣體(Parent population)性狀之各種統計方法，屬於這方面者計有各種理論分配、統計推定法、統計檢定法、變異數與互變數分析，以及新興的統計推測方法如逐次分析法、無母數統計方法、品質管制及抽

樣調查方法等，以上各種方法本書下冊亦有簡要之介紹。

本書除在內容方面較一般統計學書藉爲廣外，在編寫方面亦曾盡力做到下列幾點：第一，本書每章之末均附有各該章內容之摘要及表解，其目的在使得讀者於閱讀各該章節後，再閱讀其摘要及表解，即可收自行融會貫通之效。第二，本書對所論各種統計方法均儘量說明其來原，其目的在使得讀者不但能知其然，同時亦能知其所以然。故本書中絕大多數公式均附有證明，凡證明中涉及高深數學者，均改用文詞敘述以說明其來源，如此讀者雖未知其數學演變過程，仍能由文詞敘述中知其作成之原理及其所以然。第三，本書中一種方法如有數種不同計算公式者，均有系統地一一予以誘導而出，同時並介紹各種複雜公式之記憶方法，如此不但可使讀者易於瞭解各種統計方法之由來，同時亦可幫助讀記憶各項複雜的公式。以上數點爲著者對本書編寫的一種新的嘗試，倘能有助於讀者研習，則著者之努力並非徒勞。

本書所包含之內容較一般大學六學分統計學之教材爲多，故本書在用爲統計學之教本時，可視修習科系之不同而酌情選擇章節。當本書用爲經濟及商學等系統計學之教本時，本書上冊可逕教至簡單相關爲止，外加時間數列之相關一節及本書下冊前四章。當本書用爲生物統計及教育統計等學科之教本時，本書上冊之指數及時間數列分析兩章及時間數列之相關一節可以略去，本書下冊可教至變異數及互變數分析一章爲止。本書內容較一般書藉爲廣之原因，除爲適合不同學科之讀者及研究人士之需要外，復可使讀者在修畢初等統計學後，再閱讀其他各章，即可對古典及近代統計學之各種主要方法獲得一個扼要而基本的認識，如此可助於讀者對統計學作進一步深入研究的準備。

著者在大學時代所修之統計學爲吾師劉南溟教授所講授，著者對此一學科能發生興趣並作進一步研究，劉師諄諄善誘啓迪之功實不可

沒。著者畢業後曾隨吾師鄭堯梓教授擔任統計助教數年，在此數年中，承鄭師悉心指導，所獲教益極多。除此而外，本書在內容方面復承鄭師給予若干寶貴指示，並校閱全書一遍，衷心至為感謝。本書又承國立政治大學商學院院長韋從序教授詳細校閱一遍，並提供若干寶貴意見，對此，著者亦應致最大謝意。

著者於民國四十七年初，承國立臺灣大學農業經濟系主任吾師張德粹教授，及美國經濟文化事業協會(The Council on Economic and Cultural Affairs, New York) 路維斯博士(Dr. A. B. Lewis)之推薦，獲得該協會之獎學金，進入美國賓州大學 (The Pennsylvania State University) 之農業經濟研究所，在包倫度博士(Dr. G. E. Brandow)之指導下進修兩年，此書之成亦與此次進修有關，飲水思源，故著者對上列諸位師長所給予之協助，亦應致最深切的謝意。

本書所有插圖均由同窗唐紹東兄所繪，校對工作則由本系王嘉陵、張燕、劉蓉及何亞民等四位同學擔任，對以上諸君所給予之助力，著者亦應心感。

著者有感於近代統計學術發展之迅速，不但在理論方面是一日千里，即在應用方面亦日趨廣濶，故深感所學有限，本書謬誤之處定所難免，著者除應繼續努力更進一步外，尤望海內高明不吝賜教，以匡不逮是幸。

陳 超 塵

民國五十年十一月於臺灣大學

統計學

上 冊

目 次

第一章 緒論

一、統計學之意義與範圍.....	1
二、統計資料.....	2
三、宇宙現象.....	3
四、統計羣體.....	5
五、統計方法之體系.....	7
六、統計方法之特質與功用.....	10
七、統計方法在科學方法中之地位.....	11
八、統計學之分類與應用.....	12
九、統計學術之發展.....	13
問題一.....	16
本章摘要.....	17

第二章 統計資料之搜集

一、統計資料之種類及其搜集方法.....	18
二、原始靜態資料之調查.....	20
三、原始動態資料之登記.....	25

四、統計資料之重要性	26
問題二	27
本章摘要	28

第三章 原始資料之分類、歸類及列表

一、原始資料之整理	30
二、原始資料之分類	31
三、歸類整理之方法	35
四、統計表	41
問題三	44

第四章 統計圖

一、統計圖之意義與繪製方法	46
二、歷史曲線圖	49
三、統計地圖	53
四、長條圖	56
五、圓形圖	58
六、洛倫士曲線圖	59
七、其他重要統計圖	61
問題四	61
本章摘要	64

第五章 次數分配

一、計數、計量資料與統計分析	65
二、次數分配	66
三、累加次數分配	70

四、次數分配之型態·····	72
五、統計分析方法之次序·····	74
問題五·····	74
本章摘要·····	76

第六章 平均數

一、平均數之意義及種類·····	77
二、算術平均數·····	79
三、中位數·····	87
四、衆數·····	91
五、幾何平均數·····	94
六、調和平均數·····	97
七、各種平均數間之關係·····	100
八、各種平均數之特質與功用·····	101
問題六·····	103
本章摘要·····	104

第七章 離差

一、離差之意義及種類·····	106
二、全距·····	107
三、四分位差·····	108
四、平均差·····	109
五、標準差·····	111
六、均互差·····	117
七、相對離差·····	121
八、各種離差間之關係·····	122

八、用相關係數計算淨迴歸係數、複相關係數、估計標準誤及淨 相關係數之實例.....	342
問題十二.....	348
本章摘要.....	349

第十三章 品質相關、等級相關及時間數列之相關

一、品質相關.....	354
二、等級相關.....	366
三、時間數列之相關.....	370
問題十三.....	377
本章摘要.....	378

統 計 學

第一章 緒 論

一、統計學之意義與範圍

1-1 統計學之意義與範圍——統計學(Statistics)爲搜集、整理、分析統計資料與推論統計分析結果之科學方法，故亦有許多學者稱統計學爲統計方法(Statistical methods)。由上述定義可知，統計方法之範圍應分別如下：

- (1) 搜集統計資料。
- (2) 整理統計資料。
- (3) 分析統計資料。
- (4) 推論統計分析結果。

統計資料爲分析研究之依據，故統計工作進行之第一步即在搜集統計資料。原始資料常雜亂無章，故必須加以整理，使其成爲簡括而有系統之形式。統計資料經整理後，倘仍不能充分顯示其所蘊含之特質時，則必須進一步加以分析。如分析之資料僅屬全部研究對象之一部份樣本時，則又必須推論統計分析結果，以推知全部研究對象之可能性質。

統計方法處理之對象爲統計資料，故欲瞭解統計方法之特質及其所以然，必先瞭解統計資料之意義與特質始可。

二、統計資料 (Statistical Data)

1-2 統計資料之定義——統計資料爲在某特定時間及空間內，依據個體之特性去數計或度量社會現象或自然現象之羣體所得之資料。所謂個體 (Individual) 係泛指人、事或物的個體，凡具有共同特性之若干個體，即構成統計學上所謂羣體 (Population)。譬如每個人爲一個個體，其年齡、性別、籍貫、身長及體重等均爲人類共有之特性，全臺灣、全中國或全世界所有的人口即構成一個羣體。例如今於民國四十七年十二月三十一日午夜，將全省戶口登記簿集計得全省總人口數爲1,039,435人，其中男5,121,028人，女4,918,407人，臺北市 813,825人，基隆市 212,478人等，這些數字均符合上述定義，故爲統計資料無疑。

1-3 統計資料之三要素——由統計資料之定義及舉例可知，構成統計資料之要素有三：(1)時間，(2)空間，(3)特性。因此任何統計資料必須註明(1)時點或時間，(2)地點或地區，(3)特性之名稱及其單位，缺一不可。例如上述統計資料之時間爲民國四十七年底，地域分別爲臺灣省、臺北市或基隆市等，特性爲總人口數或不同性別人口數等。

1-4 統計資料之三特質——同時統計資料亦有三點特質：

(1) 數字性——統計資料係由數計或度量而得，故均爲數字之資料，非數字資料不得稱爲統計資料。

(2) 羣體性——統計資料係大量觀察羣體而得，一人、一事、一物既非羣體，故某一個表示單項事物特性之數字資料，亦不能稱爲統計資料。上述臺灣全省總人口數雖爲一單獨數字，但此一數字係由一千餘萬人口集計而得，故含有羣體性。

(3) 客觀性——統計資料其原始必由調查、實驗或登記而來，

故任何幻想、臆測及估計之數字亦不得稱為統計資料。

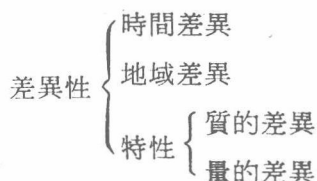
統計資料既為數計或度量宇宙間自然現象或社會現象之羣體所得之資料，故統計資料實際即為宇宙現象各個體特性之數字表示，因此統計資料所蘊含之特質必全由宇宙現象之特質而來，故欲瞭解統計方法之特質及其所以然，尚須更進一步探討宇宙現象之特質始可。

三、宇宙現象 (Universal Phenomena)

1-5 宇宙現象與統計資料——宇為縱橫四方，宙為古今上下，古今各地所有之一切現象稱之為宇宙現象。構成宇宙現象之基本因素與原理非吾人所可知，亦非吾人研究之對象，但一切宇宙現象均具有時間、空間及質和量的特性，因此構成宇宙現象之三要素亦為時間、空間及特性，而與構成統計資料之三要素相同。量可直接度量而得數字資料，質經分類數計後間接亦可得數字資料，故統計資料可以表示一切能以數字表示之宇宙現象。此即統計資料之構成必須具有時間、空間及特性三要素，以及統計方法至所以能應用廣泛之理由在此。

1-6 宇宙現象之特質——宇宙現象之基本因素與原理雖非吾人所可知，但吾人可歸納衆多之宇宙現象而獲得下列三種特質：

(1) 差異性 (Variation)——宇宙中各個體間有下列不同之差異性：



例如任兩個人，其年齡可能不同(時間差異)，籍貫不同(地域差異)，性別不同(質的差異)，以及身長、體重等不同(量的差異)等。

(2) 相關性 (Correlation)——宇宙間各因素交互影響，例如供給量影響物價，因此供給量與物價間乃有相關性之存在。所謂相關性即一般所謂之因果關係，因果事物之間必有密切的關係存在。

(3) 均衡性 (Equilibrium)——宇宙間各因素之影響雖屬雜亂無章，但却有趨向於均衡的傾向。例如影響人類體重多寡之因素有遺傳、營養及疾病等，各因素同時對體重有良好影響而使體重特高之情形，以及各因素同時對體重有惡劣影響而使體重特低之情形均較少，普通情形都是某些因素有良好的影響，另一些因素有惡劣的影響，而使得多數人有適中的體重，此種適中的體重即為多種因素影響在均衡狀態下所表現之中心數值，各人體重均有趨向於此一均衡中心之傾向。

1-7 宇宙現象之性狀——宇宙現象之變化既受上述三種特質之支配，故其結果亦有下列五種性狀：

(1) 由於各種因素有差異性，故羣體內各個體彼此間乃有差異，以致雜亂無章，使羣體所蘊含之規律無從顯現。

(2) 由於各種因素之交互影響，故一羣體內個體中之兩種或兩種以上特性間，以及兩個或兩個以上羣體間，乃有相關性之存在。

(3) 由於各種因素之變化有均衡性，故同一時間一羣體內各個體之數值乃有趨向於某一均衡中心之傾向。

(4) 由於宇宙現象各種特質綜合影響之結果，一因素隨時間變動常有週期性。例如米價由於氣候之影響，三、四月間青黃不接之時，價格較高，七、八月間新穀登場之時，價格較低，而形成以一年為週期之季節變動。

(5) 如就某一時期之靜態而言，由於宇宙現象各種特質綜合影響之結果，某一因素之羣體是有其一定之分配狀態。例如成年男子身長之分配，身材特矮及特高者人數均較少，而身材適中者人數較多，

故成年男子身長之分配常成爲中間較多兩邊較少的山狀分配。

綜合以上所述，吾人可知宇宙現象中各個體間有差異，但羣體中却含有宇宙間之通理通則，統計方法之目的即在如何由此雜亂無章之羣體中尋出其通理通則。

四、統計群體 (Statistical Population)

1-8 個體 (Individual)——統計學中所謂之個體係指人、事或物能自成一單位者，譬如一個人是一個個體，一幢房子亦是一個個體。個體之大小隨研究者之需要而定，譬如欲度量一樹之高度時，則整個一棵樹爲一個個體，又如欲度量一片樹葉之長度時，則一片葉子即是一個個體。個體之範圍愈大，則其所含之特性亦愈多，愈小則愈少。

1-9 羣體 (Population)——羣體爲具有共同特性個體之全部，譬如每個人爲一個個體，而人皆有性別、年齡及籍貫等特性，於是所有同性別、同年齡及同籍貫的人即構成一個羣體。羣體之大小隨固定特性之多寡而定，譬如人口羣體之大小即隨其固定特性之多寡而定，其情形如下：

- | | |
|-----------------|--------------|
| (1) 如毫無限制 | 所有人之羣體 |
| (2) 如限制其時間 | 現代人之羣體 |
| (3) 如再限制其地域 | 現代臺灣人之羣體 |
| (4) 如再限制其年齡(特性) | 現代臺灣20歲人之羣體 |
| (5) 如再限制其性別(特性) | 現代臺灣20歲男人之羣體 |

由上述例題看來，羣體所固定之特性愈多，則其所包含合格的個體即愈少，倘一羣體不加任何限制，則此羣體可擴大至宇宙現象之全部，但如固定羣體中個體之所有特性，則此羣體最後只剩下一個特定的個體，而不成其爲羣體了。故羣體之大小亦爲相對的，視研究者之

需要而定，譬如欲研究現在臺灣之保證糖價，則臺糖公司所有契約蔗農即構成一個羣體，其中每一蔗農即是一個個體。又如欲研究虎尾區之保證糖價，則虎尾區所有契約蔗農即構成一個羣體，該區每一蔗農為一個個體。因此所謂羣體與個體祇不過是爲了配合研究者的需要，在整個宇宙現象中所劃出之特定範圍，以及此範圍內所採用之適當單位而已。

1-10 羣體之分類——統計羣體按其本質分可分爲實在羣體及假想羣體兩種。

(1) 實在羣體——即實際存在的羣體，例如由臺灣全部人口所構成之羣體即爲實在羣體。

(2) 假想羣體——即實際不存在而假想其存在的羣體，例如從52張撲克牌中抽取5張，假想其各種可能不同的組合構成一個羣體，此即爲假想羣體。

統計羣體按其所含個體數分可分爲有限羣體(Finite population)及無限羣體(Infinite population)兩種。

(1) 有限羣體——即所含個體數爲有限之羣體，例如某年度臺灣大學全體學生所構成之羣體，因學生有定數，故爲有限羣體。

(2) 無限羣體——即所含個體數爲無限之羣體，例如一部自動製釘機之產品想像其有無限個，此無限個產品所構成之羣體即爲無限羣體。

1-11 樣本(Sample)——統計羣體如其所含個體數龐大時，研究者爲節省時間、人力及物力起見，常在羣體中抽取一部份個體加以分析，則此所抽取之部份個體對整個羣體而言是爲羣體之樣本。

1-12 羣體之性質——由前述宇宙現象之特質知，羣體內所含之特性有兩種來源，其一爲由個體之特性而來，另一即爲由宇宙間之通

理通則附着個體而來。個體之特性彼此間差異甚大，而宇宙間之通理通則則甚為穩定而有規律可尋，統計方法之目的即在尋求此宇宙間之通理通則，故統計方法必須以羣體為研究對象，而不能以個體為研究對象。又統計方法係以統計資料為其處理之對象，因此統計資料必須具羣體性之理由在此。

五、統計方法之體系

1-13 統計方法體系之由來——統計方法之目的乃在經由統計資料以分析宇宙現象所蘊含之特質或通理通則，故現今吾人至所以有某種特殊統計方法，以及這些統計方法至所以形成某種特殊體系，皆可由宇宙現象之特質方面尋得其根源。明白此種道理以後，至少有兩點好處：

(1) 瞭解統計方法之由來及其所以然後，則此種統計方法之特質及其詳細演算步驟，可由其來源而逐步推得，如此可協助吾人記憶、判斷與正確應用此種統計方法。

(2) 倘宇宙現象中某種通理通則現今尚無適當之統計方法去加以分析，或現有方法尚未能盡善盡美時，吾人乃可依該通理通則之特質以發現某種適當的分析方法。

1-14 搜集統計資料——前述統計方法之目的是在探求宇宙間之通理通則，統計資料為宇宙現象之數字表示，故統計分析工作進行之第一步即為搜集必要之統計資料。統計資料之搜集必須切合統計分析之目的，視研究對象之性質以及客觀環境等因素，而決定資料應由登記、調查或實驗而得。

1-15 整理統計資料——由於羣體內各個體雜亂無章，使羣體所蘊含之通理通則無從顯現，統計方法為欲從此綜錯複雜之羣體中尋得