

書叢行銀會學行銀

統
計
學

朱祖晦著

世界書局印行

銀行學會 銀行叢書

統計學原理

朱祖晦著

世界書局印行

統計學原理

定價二十五元

上海世界書局出版

版權所有 不准翻印

一九四九年九月初版

序

此一小冊子不足發揮統計學之精義於萬一，僅足爲統計理論入門之書籍而已。當不才最初承乏國內各大學統計學講席之時，居恆以猶爾教授之統計學入門(G.U. Yule: An Introduction to the Theory of Statistics) 爲講本。既而稍欲損益此書，俾適於經濟學系學生之用，卒未有暇也。二十二年，改任武漢大學講席，乃以三年之力編成一小冊子，於猶爾教授之書，頗有損益，且間有揚摧各家學說之處。其所主張，雖或感受裴生教授(Prof. W.M. Persons)之影響，然以不才之下愚，決難明白裴生教授之高深，總而言之，皆當視作不才之臆說而已。文稿既成，乃在武漢大學試用五年，復在其他各大學繼續試用若干年，尙無大謬，乃舉以問世。當世賢達，若進而教正之，則不才幸甚！

民國三十七年春日朱祖晦誌於鶴廬

統計學原理目錄

第一章	統計學之性質	1
	統計學之發展	1
	統計學之定義	3
	統計事實之種類	3
	統計學之內容	4
	統計研究之可靠的限界	5
第二章	統計表	7
	統計單位	7
	統計事實之分組	7
	分組之原則	8
	統計表內容之排列	11
	統計表之標題	13
	表格之畫法	13
	統計表之標目	13
	總數之位置	14
	統計事實之來源	14
第三章	普通統計圖	15
	統計圖之種類	15
	統計圖之必要條件	16
	條形圖	17
	圓形圖	21

	方塊圖統計圖畫立方體圖及統計地圖	22
第四章	曲線圖	25
	曲線圖之種類	25
	增量曲線圖	25
	時間曲線	27
	曲線之修勻	33
	四種統計曲線	34
第五章	曲線圖(續)	41
	盈縮圖	41
	累積圖	42
	迴波形曲線圖	44
	羅倫茲曲線	45
	增比曲線圖	47
第六章	比及比例率	51
	比例率之性質	51
	比例率之種類	51
	比例率之比較	54
	比例率之合併	55
	比例率之修正	57
第七章	平均數	60
	集中與代表	60
	算術平均數	60
	幾何平均數	65
	調和平均數	67

	三種平均數之關係	69
	三種平均數之特性	72
第 八 章	平均數(續)	75
	中位數	75
	衆數	80
	中位數及衆數之應用	88
	理想之平均數	89
第 九 章	離中程度	94
	各個的變異	94
	全距離	94
	四分位差	96
	平均差	97
	標準差	102
	離中係數	103
第 十 章	偏斜度及誤差	108
	曲線之區域	108
	偏斜度	111
	動矩	117
	誤差	122
第十一章	相關數	125
	統計推論	125
	相關表	125
	相關係數	127
	相關係數之顯著程度	134

	迴歸線.....	135
第十二章	相關數(續).....	138
	非直線相關.....	138
	品級相關.....	143
	虛相關.....	148
	偏相關.....	149
第十三章	指數.....	152
	指數與變動比.....	152
	比率之平均數.....	153
	總和之比率.....	155
	權數.....	156
	理想公式與因子互換試驗.....	160
	連鎖指數.....	161
	指數公式之鑒定.....	164
	固定基期與基期之移動.....	165
	全組權數與每個權數.....	166
	新增貨物.....	166
第十四章	時間數列之分析.....	168
	時間數列之性質.....	168
	季節變動.....	168
	長期趨勢.....	177
第十五章	時間數列之分析(續).....	192
	循環變動.....	192
	實業循環與相關數.....	195

統計縮減	200
第十六章 收集統計事實之方法	202
工作假設	202
從書報上收集統計事實	202
個人調查法	204
擴大的調查	205
調查區之劃分	206
調查表格之製定	207
實地調查	210
通訊調查法	210
通訊調查之表格	211
催覆函件	212
生命註冊	212
統計的推論與樣子	217
揀樣調查之可靠程度	218
揀樣調查之方法	219

統計學原理

江寧朱祖晦著

第一章

統計學之性質

統計學之發展

統計由來甚古。當埃及時代，即有人口統計，以及人民富力統計。在埃及各金字塔中，可以發現當時之人口數目，以及人民之牲畜數目。是爲統計之濫觴。基督教舊約言，所羅門 (Solomen) 王曾調查全國人口。次年，天降大疫，人民死亡者，幾及半數。由此可見，當時曾舉辦人口統計。羅馬諸帝恆以人口統計及國富統計置於座右，以爲徵兵及收稅之參考。歐洲中古時代，重商主義盛極一時。各國對於本國進出口貿易之數額，皆有統計。是時，法皇路易十四 (Louis XIV) 尤注意統計事業。對於人口，國富，公路，及進出口貿易，皆有統計。當歐洲中古時代之末年，法國拿破崙大帝亦甚注意統計事業。當時所舉辦之統計極夥。其努力有足多焉。(註)

大抵古代之視統計，僅以之爲行政上參考資料。故統計 (Statistics) 之字義，導源於拉丁文「國家」 (Status) 之義。所以表示國家之

(註) 帝王世紀詳記夏禹時代之人口數目。惜不能考證此說之由來。當蘇張遊說東周列國之時，歷舉各國之統計數字。可見當時必有統計。周禮詳述周室人口統計方法，惜此書爲偽書耳。孔子式負版者。版爲登錄人口之冊籍。蕭何入關，先收秦之圖書。所收蓋即秦之統計也。可見中國之統計事業，由來亦古。

現狀者也。當是時也，僅欲搜集許多數字，而求得綜合(Summarization)之觀念。

統計學所以卓然成爲一種科學，實由於概率理論(Theory of Probability)(註一)之發展。此爲法國數學家苟特烈(Quetelet)諸人之貢獻。

因概率觀念之發展，遂使統計學家深悟大數定律(Law of Large Numbers)(註二)。而統計學之基本觀念於焉確立。至於統計學之突飛猛進，不得不歸功於英國皇家學會(Royal Society of London)。該會中人，多思利用統計方法以研究生物學與天文學。經此提倡，統計學上各種經驗公式(Empirical Formula)(註三)日益增多。次數數列(Frequency Series)之分析方法，乃逐漸完密。最近三十年中，吾人對於時間數列(Time Series)加以深切之研究。統計學乃蔚然成立矣。於此尤應注意者，時至今日，吾人漸知統計學之效用不僅在於綜合，而在於分析(Analysis)與推論(Inference)。此乃古代統計學家

(註一)據皮爾生(Karl Pearson)之意見，概率觀念之解釋如下：

“若某事會被吾人察見，則在確屬相同之機會，吾人可以再察見之”。見 Karl Pearson: *Grammar of Science*, p. 140.

關於討論概率理論之書極多，但學者至少應讀下列二書：

(1) A. Fisher: *The Mathematical Theory of Probabilities*

從主觀觀念(A Priori concept)討論概率。

(2) J. M. Keynes: *A Treatise on Probability*

從客觀關係(Objective Relation)討論概率。

(註二)據波萊(A. L. Bowley)之意見，大數定律之解釋如下：

“以隨機揀樣法，從全體之中，揀出相當之大量事實。此等事實所含有之品性，大概常與全體之品性相同”。

見 A. L. Bowley: *Elements of Statistics*, 4th ed., pp. 278-364.

(註三)數學公式，皆從理論推演而來；故爲理論公式(Rational formula)。統計學及其他科學所用公式，皆係根據觀察之經驗而成；故爲經驗公式(Empirical Formula)。

所不及知者也(註)。

統計學之定義

一切自然界之現象，以及社會上發生之事實，皆能以數量計數或測量之。故皆得爲統計事實 (Statistical Data)。收集此等統計事實，綜合之，分析之，推論之，可以作成種種統計研究 (Statistical Treatment)。欲發表此等統計研究，或用文字，或用圖表，皆名之曰統計。

以上所述收集統計事實之方法，研究統計事實之方法，以及發表統計研究之方法，皆名之爲統計方法 (Statistical Method)。統計學者，研究統計方法之科學也。

統計事實之種類

統計事實皆能以數量計數 (Count) 或測量 (Measure) 之。然因統計學者觀察方法 (Method of Observation) 每每不同。故統計事實可以分爲兩大數列 (Series)。

第一數列之統計事實，名爲品質數列 (Series of Attributes)。此等統計事實，專門計數某種統計事實所發生之性質及其影響。例如罷工運動，共計有若干次勝利，若干次失敗。又如士兵體長，在標準以上者，爲若干人。在標準以下者，爲若干人。在研究此等統計事實之時，吾人但須考查此等性質及其影響所發生之勢力，故研究方法與通常所用者完全不同。

第二數列之統計事實，名爲變數數列 (Series of Variates)。此等統計事實，專門測量各個事實或現象變異之量 (Magnitude of Variation)。此等數列，又可分爲三大數列。第一數列，名爲彙總數列

(註)歐洲十七世紀以前，統計學之發展，應參考下列書籍：

D. Jur. V. John: *Geschichte der Statistik*, Erster Tell.

近代統計學之基本觀念，詳見下：

Karl Pearson: *Grammar of Science*.

(Categorical Series)。此等統計事實，專門測量在某一時間內，各種事實變異之量。例如工廠之中，共有童工若干，男工若干，女工若干。又如本年米之產量，甲省為若干，乙省為若干，丙省為若干。第二序列，名為次數數列 (Frequency Series)。此等統計事實，專門測量在某一時間內，某種事實所含各現象變異之量。例如兒童之體長，有一尺者，有一尺半者，有二尺者。如此等等。第三數列，名為時間數列 (Time Series)。此等統計事實，專門表示某種現象，在各個時間內，變異之量，例如病人逐日溫度。又如零售或躉售物品逐月之價格。

自實際言之，統計事實所以能分為上述四種數列，完全由於觀察方法不同。例如吾人在研究躉售價格之時，若以價值六元為標準，而觀察某種貨物所標之市價，滿六元與不滿六元者各為若干，則求得品質數列之統計事實矣。若觀察某種貨物在各市場所標之市價，則求得彙總數列之統計事實矣。若將某種貨物在同時同地所標之各種市價加以觀察，則求得次數數列之統計事實矣。若觀察某種貨物在各個時間所標之市價，則求得時間數列之統計事實矣。由此可見，一切自然界之現象，以及社會上發生之事實，其本身並非絕對屬於各種完全不同之統計數列。

統計學之內容

統計學共應研究三大問題。第一為收集統計事實，第二為研究統計事實，第三為表現統計結果。

從手續上言之，收集統計事實之方法，不外乎間接法與直接法。間接收集統計事實之法，專從書籍，報紙，公文報告，以及私家記錄上收集之。然而他人所調查者，其觀察點未必與吾人之觀察點相符合。即使偶相符合，其所調查，未必完全可靠。故間接法殊多危險。直接收集統計事實之法，或係親自調查，或係委託相當之人代為調查，或竟製就表格，交由被調查人自行報告。至於調查之範圍，若係財力與人力豐富，或有其他特殊原因，可以施行普遍的調查。若在普通情形，多

係以適當(Good)之態度,揀出若干統計事實,使之足夠(Sufficient)映現全體之品性(Characteristics)。

如何分析統計事實,此問題雖非片言可明,然其手續不外乎三種。先將已經收集之統計事實設法歸類(Classification),次乃以數學方法研究此整個數列之變異,及其勢力。如欲研究某一數列或數個數列變異之趨勢,更當以數學方法推論之。

上述二大問題,在統計學中討論至夥。然而一言以蔽之,不外乎大數定律(Law of Large Numbers)而已。所謂大數,乃顯著之數目(Significant Numbers)。此等數字,足以表現整個數列之變異。若能完全注意及之,則整個數列之品性,無不明矣。

表現統計結果之方法,不外乎圖表及其說明。此等僅為一種技能。統計學之精蘊,固不在此也。

統計研究之可靠的限界

祇知統計方法,而不知運用之,以研究一切自然科學,以及社會科學,則必不能為一健全之統計學者。然在研究進行之時,尤應深知此種研究之可靠的限界(Reliability),若不知而誤認之,則必不能了解統計學之真正立場(註)。

在開宗明義之時,吾人即應明白,統計研究祇是概括的說明全體現象(Mass Phenomena),並非說明每個份子之現象。例如根據最近學者之計算,年齡在五歲以下之中國男子,可望生存 31.0 年,年齡在五歲以下之中國女子,可望生存 32.4 年。吾人並非斷定,年在五歲以下之某一中國男子,祇可生存 31.0 年。亦非斷定,年在五歲以下之某一中國女子,祇可生存 32.4 年。不過自全體現象觀之,凡年齡在五歲以下之中國人,其生存之可能性,大概為若干年而已。此種簡括的觀念,乃統計學之重要性質。統計研究祇能說明大數之狀況,而不能說明一二個人之狀況。

(註)參考 H. L. Moore: *Law of Wages*, ch. III.

其次，吾人須明白統計研究之時間性（Time Characteristics）。所謂統計研究，不過根據某一時間所發生之現象而統計之。此等現象，未必同樣發生於其他時間。是故統計研究之可靠程度，祇限於當時，決不能移至其他時間。例如今年河南雜糧秋收甚豐。吾人決不能謂去年或明年河南之秋收，亦甚豐也。

再次，吾人須明白統計之區域性（Place Characteristics）。凡統計研究之可靠程度，祇以現今所觀察之區域為限。決不含有普遍的意義。例如今年河南雜糧秋收為若干石，吾人決不能謂今年河北雜糧秋收亦為若干石也。

最後，吾人須明白觀察之限制（Limit of Observation）。凡研究任何問題，其取材之範圍，常有一定，此種範圍，名為觀察之限制。統計研究之可靠程度，以此為限。例如研究某政黨在學生中所佔之勢力。吾人觀察之限制，僅以中學以上之學生為限。中學以下之學生，不能加入該政黨，當然不在觀察之限制以內。吾人決不可任意批評，以為合全國學生計之，該政黨在學生中勢力，並不能如此也。此等觀察之限制，本係天然發生。然吾人必須確切認識之，否則必將發生許多不公允之批評。

第二章

統計表

統計單位

欲作成統計表 (Statistic Table), 首須彙集統計事實。統計事實之單位, 名曰統計單位 (Statistical Unit)。例如某城之中, 共有男子若干名, 女子若干名。「名」即為一種統計單位。又如某省之面積, 共為若干方里。「方里」亦為一種統計單位。

統計單位必須合乎四個條件: 第一, 須簡單。第二, 須意義明白。第三, 須便於計算。第四, 須便於比較。如不能完全合乎此四個條件, 則其所規定之統計單位, 必不適用(註)。

於此有應注意者。凡在規定統計單位之時, 必須為此種單位下一確切定義。統計學家有一通病, 往往將定義不同之統計事實, 合併計算。例如各國對於息胎 (Still-birth) 所下之定義, 每每不同。通常每將各國之息胎數目, 合併於一處, 以表示全世界之息胎為若干。此種辦法, 殊不妥當。學者不可不於此三致意也。

統計事實之分組

在決定統計單位以後, 則必分門別類而清理之, 然後全體羅列清楚, 可以著手分析之矣。此等分門別類之工作, 名為分組 (Classification)。其所分成之小羣, 名為組 (Class)。每組所含統計事實之個數, 名為組次數 (Frequency)。

在分組之時, 必須注意數個問題: 第一, 所分之組必須周延。欲達

(註)參考 A.N. Whitehead: *An Introduction to the Theory of Mathematics*, ch. 17.

到此目的，必須使所分各組恰能分別包括全體統計事實，而無遺漏。例如吾人如將全國人口分爲男女兩組，則恰能包括一切。此種分組方法，卽爲周延者。然若分全國人口爲學童，壯丁，老年三組，則必不能包括嬰兒及中少年之婦女。此種分組方法，卽爲不周延者。第二，所分之組必須單純(Homogeneous)。凡同屬一組之各變數(卽統計事實)，其品性必須相同。決不可使數種品性不同之統計事實，夾雜於一組之中。例如吾人在研究工資之時，若按照工作之種類，將工人分爲數組，再將每組分爲男工，女工，及童工，則各組皆爲單純者。若但按照年齡之大小，將工人分爲若干組，則各組之品性皆爲不單純者。第三，所分之組，必須合宜(Adequate)。此事專爲研究便利著想。例如吾人如欲研究某工廠各工人工資之大小。最合宜之方法，應按照工作之種類，將工人分爲數組；例如機匠，電匠，木工，小工之類。若但曰，某號工房之平均工資爲若干，則將毫無用處。

分組之原則

自通常習慣言之，分組之原則，共有八種。第一係按照時間之先後，第二係按照區域之不同，第三係按照數值之大小，第四係按照性質之不同，第五係按照資格職分之不同，第六係按照職務或活動之不同，第七係按照功能之不同，第八係按照次序之先後。一切統計事實之分組，大抵不外乎此八種原則之一。

次數序列之統計事實，皆係按照數值之大小而分組。先將此等統計事實，各按照其數值之大小，依次排列之，成爲陣列(Array)。次再將全個序列分爲若干組。若統計事實稍多，通常所分之組，每在15組至20組之間(註)。每組所包之最小數值，名爲最低組限(Lower Class Limit)。最大數值，名爲最高組限(Upper Class Limit)。本組最高組

(註)此說亦不可拘泥。總須斟酌情形，使每組之品性皆爲單純者，且須使每組所包之變數，多數皆集中於該組之中央。則該組之中央恰爲該組最適宜之代表，分組之方法，卽無流弊矣。