

统计抽样法

邹依仁編著

新知識出版社

統計抽樣法

鄒依仁編著

新知識出版社

一九五七年·上海

內 容 提 要

本書敘述統計抽樣法的理論基礎及其具體應用。全書對抽樣法的歷史發展，大數定律與抽樣法的关系，抽樣調查的組織方式及其實際應用，抽樣誤差及抽樣單位數目的計算確定，抽樣資料的推斷等問題，都作了比較全面的說明。並對資產階級統計抽樣法作了必要的批判。書末附錄有關公式的數學證明及數表。本書可供財經科系學生及統計工作幹部作為參考進修讀物。

統 計 抽 樣 法

鄭 依 仁 編 著

*

新 知 識 出 版 社 出 版

(上海湖南路9號)

上海市書刊出版業營業許可証出015號

上海國光印刷廠印刷 新華書店上海發行所總經售

*

開本：787×1092 1/32 印張：3 3/4 字數：84,000

1957年1月第1版 1957年1月第1次印刷

印數：1—5,000本

統一書號：3076·74

定 價：(8)0.36元

目 錄

第一節	統計抽样法的概念	1
第二節	統計抽样法的史的發展	8
第三節	大数定律和統計抽样法的关系	11
第四節	抽样調查的組織方式	16
1	單純随机抽样	17
2	机械随机抽样	20
3	分層随机抽样	23
4	整群随机抽样	25
第五節	抽样誤差和抽样單位数目的确定	27
1	單純随机抽样方式的抽样誤差	30
2	机械随机抽样方式的抽样誤差	47
3	分層随机抽样方式的抽样誤差	48
4	整群随机抽样方式的抽样誤差	53
第六節	抽样資料的推断	58
1	可信任数限的推断法	59
2	直接推断法	59
第七節	統計抽样法的实际应用	61
1	農家收支抽样調查	61
2	商業統計抽样調查	77
3	工業產品質量的統計抽样檢驗	78
第八節	对資產階級統計抽样法的批判	86
附錄 1	大数定律的主要数学定理	94
附錄 2	抽样誤差公式的証明	97

附錄 3	按比例的分層随机抽样誤差和單純随机抽样 誤差的比較公式的証明	103
附錄 4	分層随机抽样誤差公式的証明	106
附錄 5	(1)不重复抽样的乘数 $\sqrt{1 - \frac{n}{N}}$ 的数值表	109
	(2)机率表	110
	(3)随机数字表	112
参考書籍		114
参考文件		114

第一節 統計抽樣法的概念

根据 1954 年 3 月苏联統計学科学會議決議，統計学是一門独立的社會科學。它在質與量的密切聯系中研究大量社會現象的數量方面，研究社會發展規律在具體地點及時間的條件下的數量表現。統計學也研究自然因素和技術因素對社會生產的影響以及社會生產的發展對社會生活的自然條件的影響。

我們知道，數理統計學是數學的一個部分，而數學是多多少少接近於作為哲學一部分的邏輯學。因此，我們絕對不能夠把數理統計學與作為社會科學的統計學相互混同起來。不過，我們亦不能否認統計學與數理統計學之間有着一定的聯系。數理統計學當然又與作為物理學一部分的物理統計學等等有着一定的聯系。

抽樣法是統計學與數理統計學間一定聯系的具體表現之一。按照抽樣法的數學理論以及抽樣公式的導演等方面來說，抽樣法是數理統計學中一個主要組成部分。但統計學在研究社會現象時，有許多時候也必須利用抽樣方法。因此，統計學就把抽樣法從數理統計學方面移用過來，並且也列為統計學中的許多專門方法之一。

統計學里面，有三個基本的專門方法：一、大量觀察法；二、分組法；三、綜合指標法。抽樣法雖然不是統計學里面的基本方法，但它與上面所述的三個基本方法都有極密切的關連。在搜集和歸納社會現象的資料時，運用抽樣法亦必須按照大量觀察和分組的原則進行的。用抽樣調查的資料推斷整個社會現象時，

主要是涉及綜合指標的研究。這就說明了抽樣法和統計學中三個基本方法的密切關連性。

在統計學里面所說的抽樣調查，實際上是指抽樣法而言的。

為了正確地說明抽樣調查的概念起見，必須先要提一提統計調查這個名詞。統計調查就是用登記統計研究綱要中所預定的各個單位的標識，科學地、有組織地搜集總體單位的原始資料。按照調查範圍的全面性與否來分，統計調查可分為全面調查與非全面調查兩大類。非全面調查尚可分為抽樣調查、典型調查、重點調查與專題論述等四種。抽樣調查、典型調查與重點調查、專題論述的主要區別是在前面兩種調查能夠把調查所得的局部資料推斷到整體，而後面兩種調查不能夠把局部資料推斷到整體。

典型調查一般是主觀地、有意識地抽選具有代表性的典型單位的。對社會現象的統計調查來說，在我國老解放區里以及全國解放後的幾年中，曾經很多次採用過典型調查方法的。從這些典型調查，也得到一些實際所需要的資料，並且解決了一些具體問題，因而，對當時的情況來說，進行這些典型調查非但是必要的，而且在調查統計工作方面，也起着某些好的作用。在我國目前業已進入社會主義建設高潮的時候，似應準備條件，逐步以科學的抽樣調查來替代典型調查了。所以，抽樣調查非但是非全面調查中最重要的一種，而且也是用途最廣泛的一種。

究竟什麼是抽樣法或抽樣調查呢？我們可以說：在一定條件下，不是抽取統計總體的全部單位，而僅僅科學地、客觀地抽取總體中一部分單位，來加以研究，其目的是獲得關於所有總體的綜合指標的正確敘述。這樣的方法，叫做抽樣法。這樣的統計調查，叫做抽樣調查。例如，對農民作收支調查時，抽取一小部分的農戶進行調查，從抽樣調查的資料，來推斷農民的收支情況、

消費農產品和工業品的數量以及勞動消耗等等。在紡織工業的生產中，抽取較小一部分的棉紗或布匹，加以檢驗，以備推斷棉紗的一般拉力和布匹的次劣品比率等等。但是我們必須着重地指出：所謂抽樣，並不是隨心所欲地亂抽一些總體單位或主觀地抽取一些單位，就可當作抽樣調查，而這種調查，必須採用有科學根據的抽樣方法，從含有大量單位的總體中，抽取為數不太多的具有代表性的單位，使能夠保證被抽取部分與整體幾乎相符合，也就是必須要使被抽取單位的特徵能夠正確地反映相適應的整體的特徵。

有關抽樣法的幾個名詞的基本概念，我們必須說明如下：

1 一般總體 我們通常把所要調查研究的事物或現象的整體，叫做一般總體。例如，我們要研究我國農家的收支情況，那末包括我國所有的農戶的總體，就是一個一般總體。

2 抽樣總體 在一般總體中，被抽取出來的一部分單位本身亦組成一個總體。這個總體，叫做抽樣總體。例如，在上述我國農家的一般總體中，用抽樣方法，抽選出一萬五千到二萬農戶來調查，這些農戶，本身便組成一個抽樣總體。

3 一般平均數 一般總體中各單位的數量標識的算術平均數，就叫做一般平均數。例如，在農家收支調查中，我們要研究全國農戶的平均收入。這種平均數，就是一般平均數。

4 抽樣平均數 從抽樣總體中各單位的數量標識所計算出來的算術平均數，就叫做抽樣平均數。例如，從一萬五千農戶的抽樣總體中，所計算出來的農戶平均收入，就是抽樣平均數。

5 一般成數 在一般總體中，具有被研究標識的這些單位數目所佔整體的比重，叫做一般成數。例如，某工廠在某一段時間內所生產的同一種產品，經過檢驗，廢品佔 2%，那末這個百分

比就是產品中廢品的一般成數。

6 抽样成数 在抽样总体中,和一般成数具有相同意义的成数,叫做抽样成数或頻度。例如,在某一种工業產品中,抽取500个產品來檢驗,發現廢品20个,那末廢品的抽样成数就是4%。

在統計学中,最常用的綜合指标,通常有三种,就是总量指标(总数或总和)、相对指标(相对数或成数)和平均指标(平均数)。在統計学中,抽样法的任务就是把从抽样調查所抽取得的資料,求出它的綜合指标,再把这些指标推断到所要研究一般总体的相应綜合指标。这些指标中,应用最广的,首推相对数(成数)和平均数两种指标。也就是說,成数和算術平均数的推断,最为广泛。由于抽样总体僅僅是一般总体中的一小部分,所以利用抽样总体的單位总数(总量指标)來推断一般总体的單位总数(总量指标)既屬不可能,而且也是不必要的。但是某些一般总体的总量指标还是可以从其他指标間接推断而求得。例如,某一地区未割谷物总收穫量的推断,虽不能从抽样总体的單位总数直接推断出來,但可从抽样总体的單位面積產量推断得全地区的單位面積產量,再把全地区的單位面積產量乘以总的播种面積,即得全地区谷物总收穫量的估計数。因此,統計学中的抽样法,对上面三种綜合指标的推断,都是用得到的。

在技術統計方面的抽样法里面,尤其是工業產品質量的抽样檢驗,除了算術平均数、成数与总数等綜合指标的推断外,还广泛地采用其他的綜合指标,如中位数、距差以及均方差等的推断。由于在生產現場中,中位数与全距可以很快地、簡單地被算出,所以这些指标的运用最为广泛。

抽样法的作用是什么呢?也就是在那种場合下,才能很好地采用抽样法?一般說來,在下列几种情況下,采用抽样法是適宜

的，而且是必要的。

(1) 当事实上不可能進行全面調查或者沒有進行全面調查的必要时，那末采用抽样法有时也能很好地完成我們的任务。在社会主义國家和人民民主國家里，对各种重要社会現象的統計資料的搜集，主要是靠統計报表制度來完成的。但是，如果每一种社会現象的資料都要靠报表或全面的一次性調查來搜集的話，那末势必使报表与核算繁复不堪。这就与核算与报表的減少和簡化的目标，恰恰相反。所以对搜集某些社会現象的資料來說，擴大抽样調查的应用範圍是具有它的重大意义的。由于这种原因，对某些社会現象的統計調查，如農民和工人、職員的收支情况等等，我們即使可以采用統計报表或者發動群众來作一次全面性的統計調查，但由于化費太多的人力、物力和財力的关系，全面調查几乎是不可能，而且亦是不必要的。也就是說，我們只有采用抽样法作为全面資料的推断根据。

对技術統計方面來說，某些資料，是不可能全面地加以調查和檢驗的，例如，工業方面某些產品，經過檢驗后就损坏或破坏了。棉紗的纖維拉力、电線、电灯泡、干電池的使用時間等等的檢驗，都是屬於破坏性的。因此，对產品質量進行破坏性的全面檢驗是不可能举行的，只有采用抽样法來了解產品的一般質量。

在工業方面，即使產品質量的檢驗不是破坏性的，但是，如果產品的价格低于產品質量的檢驗費用时，有时亦必須采用抽样法來完成檢驗質量的任务。例如，在工業大量生產过程中，用自动机床、冲床和機械化設備所制成的各种零件，它們的質量檢驗費用大都高于它們本身的价格。假定采用全面檢驗的話，勢必大大地增加產品的成本了。在这些情況下，不能不采用抽样法來弥补增加產品成本的缺陷。農產品的質量調查和檢驗，与工業品的情況几乎是相同的。

(2)采用抽样法可以作为檢查和核對全面調查的統計資料的正確性以及監督國民經濟計劃的進行。這種方法，有時亦可作為補充和修正全面調查所得基本資料的不足。例如，蘇聯在1934年7月舉行全國性的家畜普查後，再對10%的農戶進行抽样調查來核對普查的結果。我國在1953年所舉行的第一次全國人口普查，對某些地區亦曾進行抽样調查來核對這些地區的人口調查結果。由於抽样調查比較容易實施和能夠迅速地全面地展開，對加速調查過程以及順利地實現統計的基本任務，也就是對執行國民經濟計劃的過程進行經常監督是具有重要意義的。因此，在生產監督的實務中，也廣泛地運用這種抽样調查。

(3)抽样調查比較全面調查提供了按廣泛綱要對各個單位進行更全面的、更精密的、更深刻的調查、檢驗和研究的可能性。例如，對工業方面大量產品質量的全面調查和檢驗，有時限於人力和時間，不可能對每一個產品作細致的調查和檢驗，並且由於檢驗人員長期地重複工作，可能引起疲勞而發生錯差。例如，把應該合規格的產品誤作為廢品，或相反情況的錯差，亦是常常可能發生的。有時，這些錯差，可以達到相當的百分比。如果採用抽样調查，由於調查和檢驗單位數目不多，非特這些錯差的缺點可以完全避免，並且對各個個別產品可以作更精細的調查與檢驗。

列寧很重視抽样調查。他在1921年9月1日給中央統計局的信中，建議對採用集體供應的企業用抽样方法來組織詳細的研究。列寧寫道：“目前這些企業不多，還可以對它們進行全部調查……以後當這些企業很多的時候，就需用抽样方法詳細調查其五分之一或十分之一。（列寧全集，俄文版，第33卷，第13頁）

列寧在這封信中還要求對企業及蘇維埃機關的工作也組織

抽样研究。列宁着重指出：“抽选出为数不多的典型企业(工厂、国营农场)和机关——(1)最好的即模范的；(2)中等的和(3)最坏的，进行研究。(列宁全集，俄文版，第33卷，第14页)。

1954年3月苏联统计学科学会议决议中亦曾提到：目前统计学“需要最迫切进行科学研究的方法论问题中，包括关于抽样法的问题，特别是关于在居民收支调查时选择工人、职员、集体农民家庭的方法问题，……”^①

同年苏联科学院院士奥斯特洛维季扬诺夫同志在它的“关于统计学的讨论总结”中亦曾说道：“把数理统计学的方法应用到技术方面，特别是应用到产品质量的研究方面去，对此应当尽量地加以发展和鼓励，不管这是属于数理统计学方面的，或者是属于作为社会科学的统计学方面的。”

我国国家统计局发布“关于1956年全国统计工作的指示”中，亦曾说道：“重视普查而忽视抽样调查，更不愿意进行有科学根据的估计推算。……以致许多统计调查工作，费力过大，费时过长，赶不上党政领导机关的需要。”

国家统计局薛暮桥局长“在第五届全国统计工作会议上的总结报告提要”中，亦曾说道：“抽样调查是从部分推算全体，这是比较复杂的方法，需要比较高的统计知识。……现在我们还只学会了最简单的机械汇总方法。较复杂的抽样调查已开始学，还没有学会。……许多同志拒绝利用各方面的资料，以为只有自己的资料是用统计方法(机械汇总)得来的，因而也才是科学的，其他资料都是不科学的。这种思想，大大地阻碍了我们统计工作的前进。”

根据这些决议和指示，我们可以了解统计抽样法或统计抽

① 本頁的參考文件是：統計工作通訊，1954年11期，1955年2期，1956年5期中的上述有關譯文和文件。

样調查是怎樣的重要。我們应当重視這種抽样調查，好好地學習和運用它，使我們的統計工作能夠大大地前進一步。

在蘇聯統計實務中，由於抽样調查的重要性，業已廣泛地運用抽样法。除了家畜的抽样調查以外，同樣地在測定農產品收穫率，未割谷物和技術作物收穫量的預測，登記集體農莊市場的流轉額和價格以及各種不同工業部門和農業方面的產品質量檢驗，亦都採用了這種抽样法。另外，還採用抽样法來研究工人、職員和集體農民的消費和收支的調查等等。在貿易統計方面，蘇聯亦已採用了抽样法來搜集這方面的資料，並且抽样法的應用範圍，更有擴大的趨勢。我國在1954年和1955年亦已採用了抽样調查來搜集有關農民、職工收支調查的資料，而且也已獲得了一定良好的成果。

當然，由於抽样法必須在一定的抽样條件下，方始能夠很好地進行，才不致歪曲事物或現象的全面真相。因此，抽样法有它的一定用途，但有它的限制性的一面。一般說來，抽样法應用於工農業方面的產品質量和數量研究等問題，效果是最為顯著的。即使這樣，由於科學檢驗技術的進展，某些工業產品已經能夠在生產過程中，通過某些自動機器，很簡單地、自動地加以檢驗。也就是產品的制成，同時也是檢驗的完成。在這種情況下，當然就沒有採用抽样法來替代全面調查與檢驗的必要了。

第二節 統計抽样法的史的發展

俄國在17、18世紀就採用了統計抽样法，這是世界上最早的、原始形式的統計抽样法。蘇聯目前保存的有关1648年的文獻中（貴族莫羅作夫И. Морозов的企業記錄，1940年第1卷），已經運用了抽样法的原始形式。在1733年俄國出版的“御馬廐

的章程和規則”中亦已包含有抽样法的敘述。

著名的俄國統計學家豐潑洛夫(А. И. Чупров)在組織和實行抽样調查的問題的工作中具有巨大的功績。抽样法的一般理論基礎是由卓越的俄國數學巨擘契比舍夫(П. А. Чебыщев)所創立的。他對平均數極限定理的證明，為解決抽样調查的全部基本問題建立了牢固的基礎。

在1846年俄國科學院院士奧斯特洛格拉德斯基(М. В. Остроградский)便已推薦在工業產品質量檢驗中使用統計抽样法和機率的理論。他當時就提出在收進裝袋面粉和布匹(這些都是當時主要類型的產品)時，採用統計抽样檢驗法。從那時候起，復經俄國學者馬爾科夫(А. А. Марков)，利亞普諾夫(А. М. Ляпунов)以及蘇聯學者科爾麥加洛夫(А. Н. Колмогоров)、格納琴科(Б. В. Гнеденко)、羅曼諾夫斯基(В. И. Романовский)、斯米爾諾夫(Н. В. Смирнов)、雅斯德連明基(В. С. Ястремский)等改進了抽样的理論和方法。這些理論和方法目前已很有成效地被蘇聯經濟技術工作者所採用。

19世紀末葉，抽样調查在俄國地方自治局統計中獲得了廣泛的應用。

在資本主義的西歐國家里，最早採用統計抽样法的是法國大數學家萊潑萊司(Laplace)。他曾在1800年利用人口出生統計的抽样資料來估計當時全法國人口總數。以後，在1891年挪威統計局長基爾(А. N. Kiaer)舉行了有關挪威國民收入和財富的統計抽样調查。在這次抽样調查中，可能代表全國國民收入和財富情況的許多城市、鄉區和村落被挑選出來。在這些區域以內，居民按某種年齡(如17, 22, 27歲等等)、各種社會階層(五種)、各種婚姻情況以及某種字母起首姓名等標識的有關情況被詳細記錄着。從這樣的記錄來推斷出挪威全國國民收入和

財富。類似的抽樣法也由基爾和海生(E. Hanssen)应用到其他方面。例如，1894年他們应用抽樣調查來作挪威國會勞工委員會所委託舉办的年齡、收入和職業間的相关研究。這次的抽樣調查是这样進行的：在挪威首都，可能代表各階層居民的某些街道被作為挑選對象，在其他城市和鄉村亦作同樣的挑選。至于具體的抽樣是按每第十戶的成年居民的情況作詳細的記錄。最后根据這些抽查結果來推断到全國人民的年齡、收入和職業間的相关情況。抽樣法亦自然而然地推广到農業統計方面。1901年的丹麥農產品的收穫量調查就是按抽樣法來進行的。數年后，德國統計學家麥愛(P. Mayet)再采用這種抽樣法來進行德國巴登(Baden)地方的牲畜調查。

資產階級統計抽樣法的最初發展階段，虽只是反映當時實際統計工作發展的低級水平，但也奠定了這個方法的最初基礎，其中還沒有資本主義發展后期所特有的那種反動的資產階級思想。到了20世紀初，卡爾·皮愛生和它的信徒的許多著作中，資產階級統計抽樣法對於形式主義數學有了新的飛躍，實際上完全走上了唯心主義的道路。另一方面，大資本家就利用了這種抽樣法來加重工人的操作強度、加緊剝削勞動人民，為他們增加更大的利潤。此外，資產階級統計學者們更進一步地濫用統計抽樣法到各種所有的統計調查，作為偽造統計資料的假科學根據以及隱蔽階級鬥爭的重要手段之一。他們還有把統計抽樣法用來完全頂替整個統計學的意圖，來掩蓋資產階級統計學的非科學性和內容的貧乏性。這樣，使資產階級統計的道路，日趨狹隘，而漸漸走上毀滅處境了。

在社會主義制度下，統計抽樣法的任務和前途就大大地改變了。統計抽樣法在統計調查中的地位愈來愈重要了，它的应用範圍亦日漸廣泛了。但並不是象資產階級統計一樣地趨于形

式主义数学，而是真真重視这种方法的具体作用，脚踏实地地应用到各种社会现象的調查，应用到工農產品的品質檢驗，來改進这些產品的品質，減少廢品，估計总產量，为社会增加財富。这样，就使統計抽样法真正地为人民服务了。

第三節 大数定律和統計抽样法的关系

現在我們要談談大数定律以及它和統計抽样法的关系。第一个問題是：什么叫做大数定律？第二个問題是：大数定律对統計学有怎样的关系？对統計抽样法有怎样的关系？

大数定律亦称平均数定律。它是现实界的客观規律。它既可以用数学定理(主要的数学定理參看附錄1)來表現，又可以从邏輯意义、哲学意义來闡明。如果从邏輯意义、哲学意义來闡明的話，它是大量现象和过程的規律性，而且是一般只有在充分大量观察时，才会顯露现象和过程在某种具体歷史环境中具有代表性的主要特征。更詳細地說，大数定律可能象下面这样表現出來：

(1)只有在我們掌握着足够多的單位数目或足够多的情况时，大量现象的規律性及大量过程的傾向，才能很好地顯示出來。也就是說，只有在我們掌握着足够多的單位数目或足够多的情况时，对这些大量现象和过程，才能很好地進行研究。

(2)只有在平均数的形式上，这些規律性与傾向，才能被表現出來。大数定律的所以被称为平均数定律，原因就在于此。

(3)为了研究大量现象和过程，如果采取更多的單位的話，那末从这些單位的标識所計算出來的平均数或其他綜合指标，愈能够更正确地表現出这个现象或过程的規律性。

(4)如果我們研究足够多的單位數目或足够多的情况，以平均數為中心，各个單位或情况向正反兩方向的离差往往互相均衡化起來或者相互抵銷。对大量現象或过程來說，这些离差当然不是由于本質的差異所引起，而是由于偶然状态所發生的。

即使以社会現象而論，合于大数定律的例子是很多的。例如，人口統計中出生嬰孩的男女性別比例关系，在具体歷史条件下，是相当穩定的。一般是 100 个女嬰孩比 105 个男嬰孩，这是在若干年以前的統計中，已經确定了。但是这种出生男女比例規律，从个别家庭、个别时期、个别地区不一定看得出的。在个别家庭中，可能都是男孩或女孩。在个别时期、个别地区內，可能有各种各样不同比例的男孩与女孩數目。为了了解这个整个現象的正确特征，就不能只根据个别單位的調查，而只有掌握充分大量觀察，对大量事物加以綜合，才能看出这种規律。

在社会主义社会的条件下，总的來說，劳动人民的貨幣儲蓄存款总是逐步增長着的。这种增長情况，并非出于偶然性，而是具有規律性的。从这种規律性，可以反映出社会主义社会劳动人民物質福利的增加，也就是标志着人民生活的日趋改善。但是，如果就个别存戶、个别地区來看，由于某种个别原因，某一时期的存款數目可能比較前一时期不但沒有增加，而反为減少。这种由于个别原因所引起的差異是存在的，但就整个社会的發展來看，这并不否定了劳动人民的貨幣儲蓄是确实具有規律地增長着的趋势。

例如，我國解放以來，人民儲蓄增長的統計如下（見 13 頁）。

有关社会活动方面的保險事業，为了計算保險費用率以及实行人民保險公司的經濟核算制度起見，也必須了解与掌握人口死亡率和火警率等等的大数定律。当然，在社会主义社会条件下的人口死亡率和火警率是远远低于資本主义社会条件下的相