

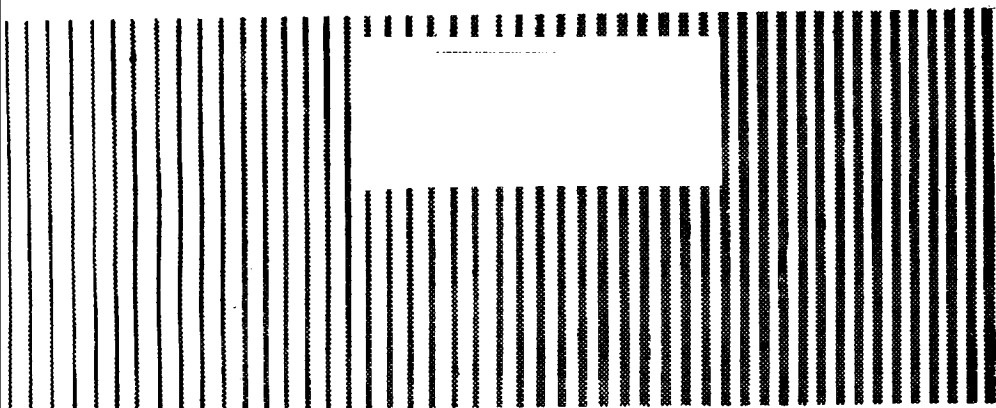
统计学原理

杨坚白 莫曰达 冯杞靖 邵祥能



统计学原理

杨坚白 莫曰达 冯杞靖 邵祥能



上海人民出版社

(沪)新登字 101 号

责任编辑 董逸伟
封面装帧 王申生

统 计 学 原 理

杨坚白 莫曰达

冯杞靖 邵祥能

上海人民出版社出版

(上海绍兴路 54 号)

本书在上海发行所发行 常熟新华印刷厂印刷

开本 850/1168 1/32 印张 22.25 插页 2 字数 509,000

1987 年 6 月第 1 版 1993 年 10 月第 5 次印刷

印数 29,001—32,000

ISBN 7-268-00195-2/F·47

定价 10.35 元

前 言

在旧中国,我们所接触到的统计学,一般都是从西方引进的资产阶级统计学者的著作;我国的学者自行写作的统计学著作或教科书,一般也是以西方的统计学为范本,其中,英美学派的统计学又是主流。

1949年,中华人民共和国成立了。从这时起,苏联的统计学源源流入我国,并且依循苏联的统计工作模式建立起我国的国家统计各级机构,开创了我国统计学界和统计工作的新局面。不言而喻,这里所说的统计,按习惯,凡是不加限制词的均是指社会经济统计而言。至于在自然科学方面,一向是运用数理统计进行观察、计量,它并不包括在上述的国家统计之内。后面将要谈的统计学上的争论,也是指社会经济统计学,数理统计工作者并未介入这个争论。

建国后输入苏联统计学之时,正是苏联展开统计学争论之际。争论的内容很复杂,大致上是围绕着两大问题而展开:一为统计学是研究规律的实质性的科学,还是研究方法论的科学?二为统计学是属于社会科学范畴的社会经济统计学,还是属于数学范畴的数理统计学?而这两大问题又每每是交错在一起的。这种争论,也随着苏联统计学的输入到我国而同时输入。其实,历史地看,这种争论并非起源于苏联。在统计学史上早就存在

实质性学派与方法论学派、社会统计学派与数理统计学派的争论。

我国在五十年代的统计学争论中，社会经济统计学派占绝对优势，原因是持不同观点者怕戴上资产阶级学术观点的帽子。不敢说老实话，于是就形成了“一边倒”。及至五十年代后期到六十年代和七十年代前半期，先之以“反右”和“四清”运动，继之以十年动乱，全国进入了学术讨论的荒漠时期。当然，在统计学上也就无所谓争论了。打倒“四人帮”后，随着实践是检验真理的唯一标准的讨论，统计学界争论再起。这一次，不再是“一边倒”，而是各抒己见，较短量长，相持不下。这本书就是在这种历史背景下，引起动机执笔写作的。

本书的宗旨是想给读者以统计学的基础知识。如书名所示，是一本讲原理的书。既然是原理，也就只限于讲述统计学的原理、原则、基本概念和基本方法，而不是学术性的专著。基于这个宗旨，我们想从正面展开论述，提供通用的统计方法。不言而喻，我们的所谓正面论述，也就意味着对自己的观点“是其所是”；其反面也就意味着“非其所非”。它既不是论战，也不是批判、评议。

为了说明本书的任务，并使读者理解本书作者的意图，先讲讲我们的基本观点是必要的。

我们认为统计学是一门广泛性、多科性的独立科学。它既是认识社会现象的工具，又是认识自然现象的工具，适用于多种现象的观察、计量和分析。在具体应用上，尽管各自选用的指标和方法有所不同，然而就其原理来说，则是共同的。即不论是对社会现象或自然现象的观察和计量，或者就统计学来说，不论社会经济统计学或数理统计学，以及各种分科的应用统计学，其基本原理和基本方法都是共同的，也就是具有共同的基础。这就

是我们的基本观点。本书的任务就是讲述这种共同性的统计学原理。在习惯上，这种共同性的统计学常常被称作通用的统计学。我们并不反对这种称呼，然而对我们来说，这种所谓通用，仅限于原理部分。

本书共计二十一章，大致上可分为五个单元。第一个单元属于总论性质，概述了统计学的性质、研究对象和统计方法论等基本问题，以及社会经济统计学和数理统计学各自的特点等。我们之所以提出两种统计学的不同特点，旨在表明，尽管这两种统计学在原理上是共同的，但毕竟是同中有异。第二个单元讲述统计工作程序，包括统计设计，统计资料的搜集、整理、分组，频数分布和统计图表等。第三个单元讲述统计基本指标，如集中指标、离散指标和相对指标等。第四个单元讲述抽样原理，统计推断和检验，包括概率、抽样原理和实践、各种推断和检验方法。第五个单元讲述统计分析，包括回归和相关、动态数列、动态分析和指数分析。考虑到统计学发展的新动向，还讲述了统计预测和统计决策。我们认为统计预测和统计决策实际上是统计分析方法的延伸，即统计分析和统计应用进一步发展的形式。本书的上述内容，可能有人认为，有些章节是属于数理统计学的范围。在我们看来，所有这些方法都是共同的，在社会经济统计中也不能不用，尤其我国现已广泛开展抽样调查，因而学习这些统计方法就更显得重要了。

本书在写作方法上，对于统计学的基本概念、基本理论和各种基本统计方法都进行了阐述和说明。在讲述统计方法需要举例说明时，一般都是既举观察社会现象方面的例子，又举观察自然现象或技术研究方面的例子。目的是为了表明这些统计方法都是共同的。但是，有必要再一次强调指出，我们所说的共同性或通用性，仅限于统计学原理及其最基本的方法，并不是说，在

统计应用上可以不分对象，不分学科，一律采用共同的方法。不仅如此，即使是可以通用的一些方法，也不是每一种方法都可以适用于任何对象。这是需要请读者加以注意的。

杨坚白 辽宁省本溪人
1911年生
中国社会科学院经济研究所研究员
中国社会科学院研究生院教授
中国统计学会顾问

莫日达 浙江省绍兴人
1926年生
国家统计局统计科学研究所副所长、研究员
中国统计学会副秘书长

冯杞靖 湖北省孝感人
1910年生
黑龙江省统计局研究员
黑龙江省统计学会顾问
黑龙江省政协委员

邵祥能 浙江省临海人
1927年生
宁夏回族自治区人民政府经济研究中心研究员
宁夏回族自治区统计学会副会长
中国数量经济学会理事、宁夏回族自治区数量经济学会副会长
宁夏回族自治区政协委员

目 录

前言	1
----	---

第一章 绪论	1
--------	---

第一节 统计学的性质和研究对象	1
第二节 统计学的发生和发展	5
第三节 统计学原理与应用统计学的差别	18
第四节 社会经济统计学和数理统计学各自的特点	21

第二章 统计方法论	29
-----------	----

第一节 大量观察法	29
第二节 综合指标法	33
第三节 大数规律	37

第三章 统计设计	44
----------	----

第一节 统计设计的意义和种类	44
第二节 统计设计的原则和内容	48
第三节 统计指标和统计指标体系的设计	52
第四节 实验的统计设计	59

第四章 统计资料的搜集	79
-------------	----

第一节	搜集统计资料的意义和任务	79
第二节	统计调查方案	80
第三节	搜集资料的方法	88
第四节	统计报表、普查、重点调查、抽样调查和 典型调查	91
第五节	调查误差	97
第五章	统计资料的整理	100
第一节	统计资料整理的概念和原则	100
第二节	统计资料整理的组织和技术	104
第三节	统计分组	109
第四节	统计表	118
第六章	频数分布	128
第一节	频数分布的概念和种类	128
第二节	组距和组数	133
第三节	组限和组中值	140
第四节	累计频数分布	143
第五节	频数分布的图示和主要类型	145
第七章	统计图	152
第一节	统计图的意义和绘制原则	152
第二节	几种主要统计图的举例和绘制方法	154
第八章	集中趋势指标	164
第一节	平均数的意义和种类	165
第二节	总体单位平均数的计算方法	168

第三节 众数、中位数、四分位数·····	183
第九章 离散趋势和偏度指标·····	194
第一节 离散趋势指标的意义和作用·····	194
第二节 离散趋势绝对数、平均数指标·····	197
第三节 离散趋势相对数·····	206
第四节 偏度指标·····	210
第十章 相对指标·····	215
第一节 相对指标的意义和种类·····	215
第二节 结构相对数·····	219
第三节 比较相对数·····	223
第四节 强度相对数·····	226
第五节 计划完成程度相对数·····	229
第十一章 概率和概率分布·····	235
第一节 概率论在统计中的应用·····	235
第二节 概率论的数理基础·····	241
第三节 随机变量及其概率分布·····	248
第四节 二项分布和普阿松分布·····	255
第五节 正态分布·····	263
第六节 选择恰当的概率分布·····	271
第十二章 抽样原理和实践·····	272
第一节 抽样原理导论·····	272
第二节 抽样法的几个基本概念·····	277
第三节 随机抽样的基本组织形式·····	280

第四节	抽样分布及有关理论	287
第五节	标准误与样本大小的关系	301
第六节	对有限总体的抽样及其校正系数	304
第十三章 统计推断和估计		308
第一节	统计推断导论	308
第二节	点估计	313
第三节	区间估计	317
第四节	从大样本计算区间估计数	323
第五节	使用 t 分布进行区间估计	328
第六节	估计中确定必要的抽样单位数	335
第十四章 统计假设检验		343
第一节	假设检验导论	343
第二节	如何检验假设	346
第三节	平均数的假设检验——大样本, 具有已知的标准差	355
第四节	测定一个假设检验的效力	359
第五节	比率的假设检验——大样本	362
第六节	平均数的假设检验——小样本	367
第七节	为平均数和比率间的差异所作的假设检验	371
第十五章 χ^2 检验和方差分析		386
第一节	χ^2 ——作为一种独立性的检验	386
第二节	χ^2 ——作为一个分布适合性的检验	396
第三节	方差分析	401
第四节	F 假设检验	408

第十六章 回归和相关分析..... 414

- 第一节 导论..... 414
 - 第二节 应用回归线进行估计..... 421
 - 第三节 相关分析..... 437
 - 第四节 围绕总体参数作出推断..... 447
 - 第五节 多元回归和相关分析..... 452
 - 第六节 防止回归和相关方法的误用..... 469
-

第十七章 动态数列和动态分析指标..... 472

- 第一节 动态数列导论..... 472
 - 第二节 动态比较指标..... 478
 - 第三节 动态平均指标..... 484
-

第十八章 动态数列分析..... 494

- 第一节 动态数列分析导论..... 494
 - 第二节 长期趋势的测定..... 498
 - 第三节 季节变动的测定..... 520
 - 第四节 循环变动的测定..... 530
-

第十九章 指数分析..... 534

- 第一节 指数的意义和种类..... 534
 - 第二节 个体指数的计算方法及其在统计分析中的作用..... 539
 - 第三节 总指数的编制方法..... 545
 - 第四节 总指数在因素分析中的运用..... 563
-

第二十章 统计预测..... 579

第一节	统计预测导论.....	579
第二节	简单模型预测.....	584
第三节	长期趋势模型预测.....	591
第四节	季节模型预测.....	609
第五节	回归模型预测.....	618
第六节	其他模型预测.....	628

第二十一章 统计决策..... 634

第一节	统计决策的基本原理.....	634
第二节	完全不确定型统计决策.....	642
第三节	先验概率统计决策.....	646
第四节	后验概率统计决策.....	651

附表..... 659

一、随机数字表	660
二、正交表	662
三、累积的二项分布表	668
四、为计算普阿松分布用的 $e^{-\lambda}$ 值表.....	685
五、标准正态概率分布面积表	686
六、t分布表.....	687
七、 χ^2 分布表	688
八、F分布表(一)	690
九、F分布表(二)	690

后记..... 693

第一章 绪论

第一节 统计学的性质和研究对象

统计学是在统计工作实践的基础上产生的；而统计学一经形成，它又对统计工作所肩负的统计资料的搜集、整理、分析等起理论指导作用。由于统计工作、统计资料和统计学三者具有如此密切的联系，所以人们习惯上把这三者通称为统计。或者说，统计一词具有统计工作、统计资料和统计学三种含义。我们在这里讨论的是统计学，当然也要联系到统计工作和统计资料。因为统计学是为统计实践服务的，而统计资料的搜集、整理和分析，又都有赖于统计学的指导。

任何一门科学，都需要首先明确它的性质和研究对象，然后才便于展开讨论和研究。可是在统计学上，这个问题一直有争论。可以说，统计学说史就是一部争论的发展史。为要讨论统计学的理论和原则，我们先从它的性质和研究对象说起，然后再讨论它的发生和发展。亦即把学说史作为分析统计学的性质和研究对象的历史源流来看待。

关于统计学的性质和研究对象的争论，核心问题是：统计学究竟是实质性科学，还是方法论科学？在历史上，两种对立的意

识形态，早已形成实质性统计学与方法论统计学的两个对立的学派。不言而喻，我国的争论也是这种争论的继续和发展。

统计学是在不断争论中发展起来的。从它诞生到现在，从来没有一个明确的、统一的定义。也就是说，直至今日也还没有一个公认的解释。凯特勒在第七届国际统计会议（1869年在荷兰召开）上说，据他的统计，统计学的定义有180种之多。美国经济统计学教授J·I·葛里芬在他所著《统计方法与应用》一书的序言中不无感慨地说：“统计的应用范围如此之广，以致很难对这门学科下个定义。”还有人形象地说，有多少统计学家就有多少种定义，即使是最精明的统计学家也统计不清。当要提出我们对统计学的见解之前，首先指出这一情况，旨在表明我们的看法，不过是百家争鸣中的一家之言而已，见仁见智，谁是谁非，自当由历史实践来作检验。

我们认为统计学是研究社会现象和自然现象关于总体数量关系计量的统计方法和统计方法论的科学。它以大量观察为基础，以平均数为中心，通过统计总体，为揭示数量关系的大数规律提供一系列的计量、分析方法。

下面，对我们的这个看法作几点说明。

首先，我们在习惯上，当谈到统计、统计学而不加限制词时，一般是指社会经济统计、社会经济统计学。然而严格地说，这种说法在今天已经不够确切了。因为统计和统计学所观察、研究的对象并不限于社会现象，也包括自然现象。有人也许会用列宁的话来反驳说，列宁不是说过统计是认识社会的有力武器吗？^①怎能将自然现象统计也拉进来呢？诚然，列宁是说过这个意思的话，然而他并没有说，统计不能作为认识自然现象的武器。何

^① “社会认识的最有力的武器之一的社会经济统计”这句话，列宁是用以批判为统计而统计的德国资产阶级统计。参阅《列宁全集》第16卷，第431页。

况历史实践早已表明，统计和统计学是既用于社会现象又用于自然现象的观察和研究。所以，我们应该把社会现象与自然现象并提。

其次，我们肯定了统计学是研究数量关系的方法论的科学，而不是研究实质性问题的科学。这里面已经包含了研究对象，即它所研究的客体是社会现象和自然现象在总体上的数量关系，它是通过自己的统计方法和统计方法论的研究而服务于数量关系的研究的。这也就是统计科学在社会领域和自然领域中所要研究解决的特有的矛盾。众所周知，除了统计学之外，不再有其他科学来研究这种方法和方法论。

第三，方法论科学是否也探讨规律性呢？我们认为从哲学上说，辩证唯物主义者是把方法论与认识论看作是一致的。本质性的认识就是认识事物的规律性。统计学研究和提供一系列的统计方法，其目的就在于认识并且揭示数量关系的大数规律。正是在这个意义上，有的人把大数规律叫作统计规律。在这一点上，统计学也如同数学一样，是通过自己的公式和方法等等来揭示数量变化的规律。所不同的是：数学是揭示一般的数量规律；统计学则是揭示一定的统计总体的数量变化规律，而总体的同质性是统计学的一个最基本原则。

第四，既然统计学也揭示规律性，为什么不把它称作是反映规律的实质性科学呢？请注意，我们在这里说的规律，仅仅是指在一定统计总体下的数量变化的规律，而不是社会经济领域和自然领域千千万万事物本身的各个实质性规律。当然，它也是实质性规律的数量表现。我们在前面已经指出过，统计学是一门广泛性、多科性的科学，它不可能把大千世界中各有千秋的、千差万别的客观实体的具体规律统统在一门统计科学中表达出来。然而不同事物在数量上却有共同性的规律，它是可以通过