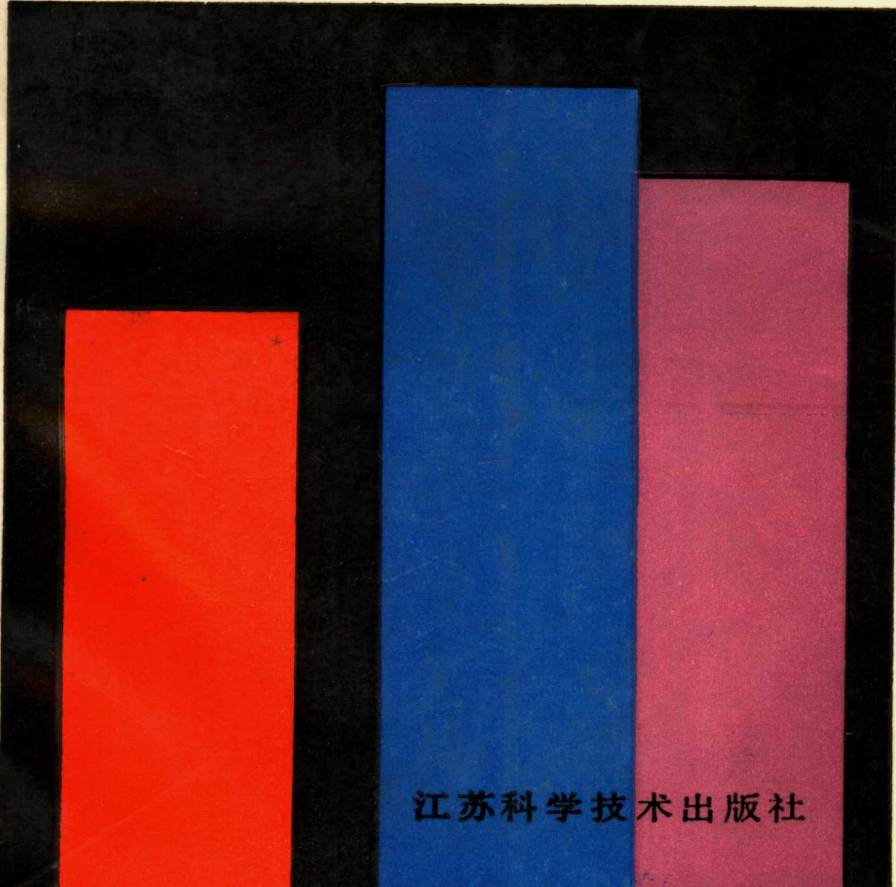


周 铭 编著

统计学原理

TONGJIXUEYUANLI



江苏科学技术出版社

统 计 学 原 理

周 铭 编著

江 苏 科 学 技 术 出 版 社

统计学书系

基础统计

统计学原理

周 铭 编著

出版：江苏科学技术出版社

发行：江苏省新华书店

印刷：南京人民印刷厂

开本787×1092毫米 1/32 印张14.625 字数321,000

1987年9月第1版 1987年9月第1次印刷

印数1—11,000册

ISBN 7-5345-0154-7/F·22

统一书号：4196·040 定价：3.10元

责任编辑 周兴安

前 言

当前，我国各个方面都在进行改革，试问统计教材要不要改革？回答应该是肯定的。50年代初期，我们全盘引进了苏联的统计学，认为统计学是一门阶级性很强的社会科学，自此以后，学术问题就蒙上了政治色彩，从而在这门科学领域里形成了一花独放的局面，苏式统计教材统治了我国讲台近30年。近年来，统计学界普遍感到这种数十年一贯制的教材不但体系刻板，而且内容也贫乏，远远不能适应统计科学在我国的发展与四化建设的客观需要，因而有改革的必要。但迄今为止，很多统计教科书和讲义虽然增添了一些数理统计方法，但对原有体系和内容则几乎没有什么触动，两者杂处类似拼盘，不很协调。看来，对传统的部分要作较大改革是颇费力气的。本教材试图在这方面迈出一大步，其方向是否正确，成效究有多大，尚有待于实践检验。

统计学是一门统一的科学，还是如当前统计学界多数人主张的分属于社会科学与自然科学的两门性质不同的科学，当然还需要继续探讨。但是，统计学所包含的内容如此之丰富，想在一本著作里兼收并蓄是不可能也不必要的。不同的统计著作或教材有不同的重点，这是客观的必然。因此，“一门”、“两门”之争，并不妨碍或许还能促进统计教材的多样化。况且，有些统计方法主要适用于对某类对象的研究。例如，动态分析与指数分析就是适应研究社会经济现象的需要而产生的，而方差分析与正交试验则主要用于实验性的科

学研究。因此,统计教材的内容应针对具体对象而有所侧重。本教材的对象是经济类专业的学生。

本书与其他多数教材相比,除了在体系上的某些区别(如抛弃了“统计工作阶段”的模式)外,还压缩或避免了一般常识性的和流于空泛的论述,增加了一些在以往教学与实践我感到有用的内容,并佐以丰富的例题。

习惯上将在概率论基础上发展而成的统计方法称为数理统计方法(国外很多地方称之为推断统计)。根据需要,本书也和大多数教材一样,采用了其中抽样估计与相关分析两部分。如果已学过或即将学习数理统计学,则此两章可略去。

考虑到一般学校均不单独开设经济预测课程,因此列入了统计预测一章。

中等专业学校若采用本教材,可根据实际需要自行调节深度与广度,或省略某些章节。

由于我系蒋琳同志对动态分析的某些内容进行过研究,有新的见解,因此请他撰写了第六章初稿。

本书的一些重要章节曾请西安统计学院王广森教授与我系汪荫元教授作了审阅,南京粮食经济学院顾伟林副教授不仅仔细校阅了第九章内容,并对该章的例题用微机进行了复核,我系潘文珠老师亦对本书提过宝贵建议,谨在此对他们致以衷心的感谢。

作者水平有限,成书仓促,特别是有些观点尚属探索性质,因此错误、疏漏在所难免,殷切期待广大读者批评指正。

作者

1987年2月于南京农业大学农业经济系

目 录

第一章 导论	1
§ 1 统计学的科学性质	1
§ 2 统计学中的几个基本概念	3
总体与个体 标志 变量 指标	
§ 3 原始资料的搜集	7
统计调查的意义 原始资料的来源 统计调查的方式	
第二章 原始资料的归纳整理	11
§ 1 统计分组	11
原始资料归纳整理的内容 统计分组的作用	
分组标志与分组界限 简单分组与复合分组	
§ 2 分布数列	19
分布数列的概念 分布数列的种类 分布数列区间的表示 等距数列组距与组限的确定 变量值的归组	
§ 3 绝对数	26
绝对数的概念 总量指标的重要性 时期总量与时点总量 总量指标的计量单位	
§ 4 统计表	28
统计表的意义与重要性 统计表的结构与编排 编制统计表的一般规则	
§ 5 统计图	32
统计图的意义与重要性 统计图的种类和用途 绘制统计图的一般规则	
第三章 相对数	38
§ 1 相对数的概念	38
统计分析的基础方法 对比分析 相对数的概念与表现形式 相对数的作用	

§ 2	比重相对数	43
	比重相对数的概念 比重相对数的表现形式	
§ 3	比较相对数	44
	比较相对数的概念 比较相对数的表现形式 差额相对数	
§ 4	计划完成相对数	47
	计划完成相对数的概念 计划进度执行情况的检查 长期计划执行情况的检查	
§ 5	相对数的计算与应用	50
	相对数的计算 两个相对数的对比分析 相对数的应用	
第四章	平均数	60
§ 1	平均数的概念	60
	变量的代表值 平均数的含义与两类平均数 平均数的作用	
§ 2	等分性平均数	62
	等分性平均数的概念 等分性平均数的计量单位	
§ 3	算术平均数	64
	算术平均数的概念与基本公式 算术平均数与等分性平均数的关系 加权与权数 从部分平均数计算总平均数 从分布数列计算平均数 比重权数 派生变量的平均数 算术平均数的数学性质与简算法 0—1 变量的平均数	
§ 4	倒数平均数、几何平均数与平方平均数	83
	倒数平均数 几何平均数与平方平均数	
§ 5	众数与中位数	91
	位置代表值 众数 中位数 众数、中位数与算术平均数之间的关系	
§ 6	平均数的计算与应用	98

平均数的计算 平均数的应用

第五章 变异数	104
§ 1 变异数的概念	104
变量的离散值 变异数的概念与作用	
§ 2 极差与平均差	105
极差 平均差	
§ 3 标准差与方差	107
标准差与方差的概念 从分布数列计算标准差 方差	
的数学性质 标准差的简算法 0—1 变量的标准差	
标准差系数	
§ 4 方差分解定理	119
组方差的代表值 方差分解定理	
第六章 动态分析	123
§ 1 动态数列	123
动态数列的意义 动态数列的种类 动态数列的编制	
原则	
§ 2 动态分析指标	126
动态分析指标的意义和种类 动态数列的两两对比分	
析 动态数列的代表值 动态分析指标数列的代表值	
§ 3 长期趋势的测定	144
长期趋势的概念 移动平均修匀法 趋势线配合法	
§ 4 季节变动的测定	166
季节变动的概念 简单同期平均法 趋势剔除法	
第七章 指数分析	175
§ 1 指数分析的任务	175
指数的概念 指数分析的任务	
§ 2 总指数计算方法的研究	177
总指数的性质 总指数的两种计算方法	
§ 3 综合指数	179

同度量问题 同度量因素的确定 同度量因素的固定 时期 拉氏指数与帕氏指数	
§ 4 平均指数	187
个体指数的加权 权数的同度量 平均指数与综合指 数的关系	
§ 5 个体指数的两因素分析	194
因素分析的概念 因素分析的基本方法 因素分析中 相对变动与绝对变动不成比例 量变与质变的顺序	
§ 6 总指数的两因素分析	201
指数体系 拉氏指数与帕氏指数的搭配 指数体系在 推算上的应用	
§ 7 总平均数指数的两因素分析	212
总平均数指数构成的剖析 总平均数指数的指数体系	
§ 8 多因素分析	218
个体指数的多因素分析 总指数的多因素分析 总平 均数指数的多因素分析	
§ 9 宏观指数	231
宏观指数的概念 宏观物价指数 宏观产量指数 指 数数列	
第八章 抽样估计	235
§ 1 抽样估计的一些基本概念	235
抽样估计的意义与重要性 样本与样品 抽样误差 抽样的随机原则 重复抽样与不重复抽样 置信区间 与置信概率	
§ 2 概率与随机变量	238
随机事件与概率 古典概率的计算 概率的两个重要 运算定理 随机变量 概率分布 概率密度函数 期 望与方差	
§ 3 几种基本分布	247

二项分布 超几何分布 正态分布 中心极限定理

§ 4	样本统计量.....	253
	样本数据的性质与特点 样本平均数的分布 样本平均数的期望 样本平均数的方差 标准误 样本成数的分布与特征数	
§ 5	总体平均数与总体成数的抽样估计.....	260
	总体平均数的抽样估计 总体成数的抽样估计 小样本的应用 相对误差限	
§ 6	必要样本容量.....	272
	必要样本容量的意义 估计 $\bar{X}$ 的必要样本容量 估计 P 的必要样本容量	
§ 7	抽样的设计方式.....	277
	纯随机抽样 等距抽样 分组抽样 整群抽样 分级抽样	
第九章	相关分析	292
§ 1	相关分析的性质与任务.....	292
	相关的概念 相关的种类 相关的实质 相关分析的任务 相关性质的判定	
§ 2	相关直线.....	298
	线性相关的形成 相关直线的意义 相关直线的特点 相关直线的配合	
§ 3	从变量的区间估计.....	310
	相关估计中的标准误 对条件值平均数的区间估计 对条件值的区间估计 置信带	
§ 4	相关系数.....	318
	相关系数的意义 相关系数的基本公式 相关系数的积差公式 对总体相关系数的估计 等级相关法	
§ 5	双项分布表.....	334
	双项分布表的意义与编制 双项分布表对计算的简化	

§ 6 多元相关分析.....	340
多元相关分析的概念 二元相关直线 复相关系数与 净相关系数	
§ 7 一元非线性相关分析.....	355
非线性相关的概念 指数曲线相关 二次曲线相关 双曲线相关 曲线类型的比较选择	
第十章 统计预测	375
§ 1 统计预测的一般问题.....	375
统计预测的概念 统计预测的任务与实质 统计预测 的种类 统计预测的误差 统计预测的基本方法 配 合预测模型的一个原则 统计预测中权数的形式	
§ 2 加权最小平方预测法.....	379
线性趋势的预测模型 指数曲线趋势的预测模型	
§ 3 指数修匀预测法.....	385
一次指数修匀法 二次指数修匀法 初始值的估计 修匀系数的取值 三次指数修匀法	
§ 4 其他配合预测方程的方法.....	400
分段加权平均法 自身相关预测法 朴素模型预测法	
§ 5 相关分析预测法.....	410
§ 6 季节模型预测法.....	415
思考题(附解答或提示)	423
附 录	445

第一章 导 论

§ 1 统计学的科学性质

从80年代起，我国的统计科学逐渐改变了过去僵化的面貌，得到令人兴奋的发展。这主要是由于解放了思想，基本上消除了某些偏见，从而大大地丰富了统计学的内容。

经过这几年的争鸣，目前统计学界对于统计学并非实质性科学而是方法论这一点已很少有异议。但是，在统计学的科学性质方面仍然存在分歧，其根本的不同意见在于：究竟统计学是一门统一的科学，还是存在着两门性质不同的统计学——属于社会科学的社会经济统计学与属于自然科学的数理统计学？由此而引起了有关统计学的对象和理论基础等看法的相悖。

比较流行的看法是：并存着两门性质不同的独立的统计学，即社会经济统计学与数理统计学，它们有各自不同的研究对象和各自不同的研究内容、研究方法。数理统计学是应用数学的一门分支，属于自然科学。社会经济统计学属于社会科学的范畴，“它的研究对象是大量社会经济现象的数量方面，……以认识社会发展规律的具体表现”或“社会经济统计学属于认识社会经济现象数量方面的方法论科学”。这门科学有自己的理论基础，“社会经济统计学以历史唯物主义和马克思列宁主义政治经济学为理论基础”或“社会经济统计学以社

会经济现象为其研究对象，这就规定了它必然以社会经济理论科学为其理论指导，……社会经济统计在研究社会经济现象时，就必须以政治经济学所阐明的范畴和规律为依据”。（以上引自《社会经济统计学原理》，高等学校文科教材；《社会经济统计学原理教科书》，高等学校文科试用教材）。

本书作者认为：统计学不是一般的社会科学或自然科学，它为其他科学提供从量的方面着手研究客观事物的方法和理论。统计方法是适应各门科学的客观需要而产生的，因此内容广泛，形式多样，有以一般数学理论为基础发展起来的方法，也有以概率论为基础发展起来的方法。它们多数可以应用于一切领域，也有的主要用于某些领域。不同的统计著作当然可以对之分别或有重点地论述，但这些内容都是统计方法，都属于统计这一门科学。从科学性质上来说，并无将统计学归入哪一类科学的必要。如果客观上有需要非归类不可，则由于当前我国在社会经济领域里运用统计方法更为广泛和迫切，发展得更快，因此将之隶属于社会科学也没有多大关系，这并不会改变统计学的方法论性质和影响它在自然与技术领域中的应用和发展。

统计学的研究对象是什么呢？它不是研究具体的社会现象和自然现象的什么具体规律，否则它就成为一门实质性的科学了。它只提供研究这些现象的方法。如果要问它自己研究什么，可以这样说：它研究如何通过事物在具体时间、地点条件下的数量表现，来分析事物的特点和规律。因此，它研究的也就是方法本身。

作为一门科学的统计学与统计部门和有关部门实际进行的统计工作不能等同起来。作为一门方法论的科学，统计学研究的只是统计方法本身，其最终目的在于使自己成为其他

有关科学和实际工作者认识客观现象的强大武器，因此除了一切科学必须遵循的唯物辩证法和统计方法本身所必需的数学基础外，它无需再以其他社会科学或自然科学作为自己的理论基础。在实际进行统计工作时，必然要有自己的具体对象和具体目的。因而，必须在有关具体对象的科学理论的指导下，才能有效地运用统计这一武器。这就是说，在研究社会经济现象时，必须以政治经济学与其他有关社会科学为理论基础；在研究自然现象时，必须以有关的自然科学为理论基础，这是毋庸置疑的。

§ 2 统计学中的几个基本概念

总体与个体 统计之所以成为研究和认识事物客观规律的有力武器，主要原因之一在于它对事物进行的是大量观察，即观察事物的群体，而不是只注意个别现象。人们很多错误观点的产生常常是因为其只根据个别现象就作结论。例如，“我祖父一辈子不刷牙，牙齿仍然很结实，刷牙岂非多此一举”；“这块肉虽有些变质，我吃了不是没事吗？还是俗话说得对，不干不净，吃了没病”；由于气象台一二次预报不准就否定气象预报的科学性；根据个别农户这两年生产不景气而否定生产责任制的优越性；只掌握一二个数据就作为科学研究的成果等等。以上这些论断和做法显然不能使人同意，它们正如列宁所批评的那样：“胡乱抽出一些个别事实和玩弄事例”。如果对上述问题进行大量观察，结论就会改变。而坚持进行大量观察正是统计学的基本观点。

用统计方法作具体研究时，被研究事物的群体称为总体，组成总体的单位则称为个体（或总体单位）。例如，当

我们调查某地区轻工企业的经营管理情况时，该地区的全部轻工企业就组成一个总体，而各个轻工企业就是个体；当研究某企业职工的文化状况时，该企业全体职工就是总体，每一职工就是个体。

可见，总体总是同类事物的某种范围的集合，个体则是此集合的元素。

总体与个体是灵活的概念，可因研究范围、目标和研究的阶段而变。例如，研究某市区居民家庭经济收入时，该区所有的住户构成总体，而个体是户；如果同时又研究该区居民的年龄组成时，全体居民组成总体（此时总体包括的范围与前述并无区别），而个体则是人（组成总体的实体与前不同）。又如，在我们专门研究个体的某一项数据时，总体就转化成了数据的集合，而个体则是每个数据。此时，这群数据在数值大小上的分布情况就称为总体的分布，这群数据的平均数就称为总体的平均数。

总体包含的个体数一般总是大量的和已知的，但有时量极大而不可知，甚至理论上能达到无限。例如，某个品种的家畜（它应包括现有的、过去饲养过的和今后繁殖的）、某种型号的产品（它也应包括已经生产过的和今后可能继续生产的）、某项科学试验的数据（人们可以永远不断地重复作这项试验）等，我们将这一类总体称为无限总体。对这一类总体的研究，只能采取特定的手段——统计抽样法。

标志 我们是通过个体来研究总体的，而对个体的观察，则立足于个体的种种特征。例如，个体为企业时，它的所有制、隶属关系、规模大小、产品产量、成本、劳动生产率等，都是彼此有所不同的特征；个体为人时，它的性别、文化程度、民族、年龄、身高、体重等，也都是相互区别的特征。在统计学中将

个体的这些特征称为标志，它是观察个体的着眼点。可以说标志就是对个体所作的某种描述。

标志分为两类。如果它的具体表现是一个量（象上面举的企业的产品产量、成本，人的年龄、身高等），则称之为数量标志；如果它不表现为量（象上面举的企业的所有制、隶属，人的性别、文化程度等），则称之为属性标志（或品质标志）。有些属性标志可以表现为数字，如产品质量可以分为一、二、三级，通讯地址可以表现为门牌号数。其他如人的出生年月、运动员的号码等均是。但这些数字并不代表量，不能对之作数学运算，它们只代表一个顺序，其实质与用甲、乙、丙或A、B、C之类来表示并无区别。因此不应将这类属性标志误认为数量标志。

在具体的统计分析过程中，数量标志与属性标志也可以互相转化。数量标志可转化为具有某一指定数量区间的属性标志，例如，根据需要，可将40至50岁的知识分子定为中年知识分子；将小麦亩产超过400公斤的田块定为高产田；将使用寿命达不到指定时期的产品定为次品等。反之，根据需要，可将属性标志转化为只取0与1两个值的数量标志。

变量 数量标志在不同的个体上表现为不同的量，因此数量标志又称为变量。当具体观察一个个体时，该变量就表现为一个具体的值，如某产品的产量为“10万吨”、某人的年龄为“35岁”等。变量的这些具体数据称为变量值（或标志值）。

按照变量值的表现形式可将变量分为两类。如果各种可能的变量值之间是间断的（在数轴上表现为分散的点），因而是可以一一列举的，就将这种变量称为离散变量或不连续变量。象职工人数、牲畜头数、商业网点个数等这一类由计数所得而只能取整数值的变量都是离散变量。而一些商品（如鞋

帽、油漆等)各种规格的长度、重量、容量之类变量,虽能取小数,但数与数间是间断的,因而仍属于离散变量。如果各种可能的变量值之间是连续的(在数轴上表现为紧密相连而无空隙的点),因此在一定区间有无限多个而不可能列举,就将这种变量称为连续变量。连续变量的值一般均由测度而来,如时间、长度、重量、体积等,它们的值理论上可以持续到无限位小数。对这两类变量的进一步加工分析方法有所不同。

按照变量值的性质,也可将变量分为两类。如果变量值是个体的固有数值(如个体为职工时,其年龄、工资等;个体为家庭时,其成员数、消费支出等;个体为企业时,其职工人数、总产值等),则称为基本变量或原始变量;如果变量值是个体上其他固有值的某种加工结果(如个体为职工时,其平均每月出勤天数、所生产产品的次品率等;个体为家庭时,其消费支出比重、平均每人生活费用等;个体为企业时,其全员劳动生产率、产值计划完成百分比等),则称为派生变量或次级变量。对这两类变量的进一步加工分析方法也有所不同。

指标 指标这个概念已很普及,在计划工作中或其他方面是常采用的。但在统计中,指标则有其特定的意义:由原始资料(指个体的各项标志的具体表现)整理加工所得的数量结果称为指标。有时为与其他方面相区别,称之为统计指标或综合指标,指标可以表明被研究对象(指总体或它的一个部分)的某个方面的情况、特点或规律。例如,根据某企业每个工人的性别、年龄、工种、月工资额等资料,经过综合加工,可得诸如工人总数、女工比重、工人平均年龄、某工种工人的平均月工资额和工资额的差异程度、工人比上年增加的数量和增长的速度等等,这些都是指标。可见,指标是采用统计手段计算而来的,统计方法所取得的成果正是通过种种指标来