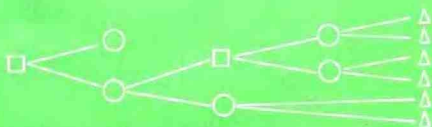


实用统计学

主 编 关长铭 王红蔚
副主编 程世辉 郭斌彩



列号	1	2	3	4	5	6	7
因素	A	B	$A \times B$	C	$A \times C$	$B \times C$	



成都科技大学出版社

97
08
26

实用统计学

主 编 关长铭 王红蔚
副主编 程世辉 郭斌彩



成都科技大学出版社

(川)新登字 015 号

责任编辑 王泽彬

封面设计 郭改文

实用统计学

主 编 关长铭 王红蔚

副主编 程世辉 郭斌彩

成都科技大学出版社出版发行

河南省郑州市宏新印刷厂印刷

850×1168 1/32 12.25 印张

1996 年 10 月第 1 版 1996 年 10 月第 1 次印刷

印 数:1—1000 册 字数:292 千字

ISBN7-5616-3295-9/F·340

定价:17.80 元

前 言

随着科技进步与社会发展,人们要认识的客观世界更加广泛与复杂,统计方法作为定量分析的一种重要工具,已成为各个领域不可缺少的手段而日益受到重视。在今天,您就是现代社会的一个普通居民,不懂些统计知识也恐怕难以应付大量的数据信息。特别地,电子计算机的发展与普及,使得经常地从大量复杂的数据中提取信息成为可能。“有人认为,统计知识还将进一步普及,它一定会比解二次方程更为人所知,这恐怕不是言过其实的。”

自从数理统计学成为统计学一部分以来,统计学就有“描述统计学”与“推断统计学”之分。描述统计学研究的是如何搜集统计资料,以及将搜集到的资料用图、表以及一些概括性的数字加以表示的方法,也包括一些曲线的拟合方法等。它使统计资料更加简明扼要,更便于认识客观事物。它是统计学的基础。推断统计学则是根据样本信息去推断总体(包括总体分布、参数、独立性,以及预测、控制等等)。由于实际问题的条件限制,我们往往只能作部分观测,要达到对总体的全面认识,就必须依赖于推断。因此,推断统计学就成了统计学的核心。奉献给读者们的《实用统计学》一书没有侧重于哪一部分,而是试图从描述统计与推断统计相结合上较全面地来介绍统计方法。

统计学又有理论统计学与应用统计学之分。前者是研究统计方法的数学基础的,例如各种估计原理、假设检验原理、统计决策原理等等。后者则是把统计方法应用于各个领域而形成的,例如社会经济统计学、生物统计学、教育统计学、工程统计学等。由于各个领域要解决的问题既有共性又有特殊性,所以各专业统计学中的统计方法也就既有共性又有特殊性。《实用统计学》不研究统计方法的数学基础,虽然有些地方也作些数学推导,但那仅仅是为了说

明方法的合理性或可靠性,数学基础相对薄弱者可以略去不读。《实用统计学》也不是为解决某一领域里的问题的专业统计学。本书以“通用性”、“实用性”为原则选择并介绍常用的统计方法。多元统计方法也有很强的实用性,但由于层次较高且这里篇幅有限,所以就割爱了。另外,鉴于市场经济和教育工作的需要,本书较多地选取了经济管理及教育研究方面的方法和例子,这也算本书的一种倾向吧!

本书各章有一定的联系,但也有相对独立性。读者可以根据需要进行选择,例如把第一、二、三、四章及第八、九章放在一起,就可以作为社会经济统计学来读;五、六、七三章就是一般的数理统计学;仅读第十章就可以基本上学会试验设计方法;只读第十一章就可以基本上学会统计决策方法等。

会使用电视机的人未必懂得电视机原理,不懂得统计学数学基础的人一样可以学会使用统计方法。我们希望《实用统计学》能帮助数学基础相对不足而又需要使用统计方法者,尽快学会其所需要的统计方法。

本书由关长铭、王红蔚主编,程世辉、郭斌彩任副主编,安春香、高景丽、参加了部分内容的编写。具体分工是:第一、二、十一章由关长铭(河南教育学院)执笔,第五、七章由王红蔚(河南教育学院)执笔,第八、九章由程世辉(河南教育学院)执笔,第十章由郭斌彩(商丘教育学院)执笔,第三、六章由安春香(焦作教育学院)执笔,第四章由高景丽(河南机械工业学校)执笔。全书由关长铭完成统审定稿工作。

在编写过程中参阅了有关著作,引用了一些例子,恕不一一指明出处。在此一并向有关作者致意!

由于时间仓促及水平有限,书中缺点或错误定所难免,恳请专家和读者批评指正!

编著者

1995年10月

目 录

第一章 总论

- 第一节 统计学的研究对象····· (1)
- 第二节 统计学的学科体系及研究方法····· (6)
- 第三节 统计学的基本概念····· (9)

第二章 统计设计

- 第一节 统计设计的概念、种类和基本内容····· (16)
- 第二节 指标与指标体系设计 ····· (19)
- 第三节 统计整理方案的设计 ····· (21)
- 第四节 统计调查方案的设计 ····· (26)

第三章 统计调查与整理

- 第一节 统计调查与整理的意义 ····· (30)
- 第二节 统计调查 ····· (32)
- 第三节 统计整理 ····· (46)

第四章 统计特征数

- 第一节 集中量数 ····· (60)
- 第二节 相对地位量数 ····· (82)
- 第三节 统计变异数 ····· (86)
- 第四节 偏度与峰度 ····· (97)
- 第五节 统计相对数(相对指标)····· (101)

第五章 统计推断

- 第一节 抽样分布····· (107)
- 第二节 参数的点估计····· (111)
- 第三节 参数假设检验及区间估计····· (117)
- 第四节 抽样误差及样本容量····· (138)

第五节 非参数假设检验·····	(150)
第六章 方差分析	
第一节 单因素方差分析·····	(167)
第二节 方差齐性检验及多个均值的逐对比较·····	(175)
第三节 两因素方差分析·····	(179)
第七章 相关与回归分析	
第一节 相关分析·····	(189)
第二节 一元线性回归分析·····	(193)
第三节 一元非线性回归分析·····	(206)
第四节 多元线性回归分析·····	(214)
第八章 时间数列分析	
第一节 时间数列的意义和分类·····	(221)
第二节 时间数列分析的主要指标·····	(226)
第三节 长期趋势的分析·····	(238)
第九章 统计指数分析	
第一节 统计指数的概念、种类和作用 ·····	(251)
第二节 总指数的编制·····	(253)
第三节 指数体系和因素分析·····	(262)
第十章 试验设计法	
第一节 单因子试验·····	(278)
第二节 双因子试验·····	(291)
第三节 正交试验设计法·····	(302)
第十一章 统计决策方法	
第一节 决策的含义及分类·····	(322)
第二节 风险型决策·····	(327)
第三节 不确定情况下的决策·····	(351)
第四节 贝叶斯决策·····	(355)
第五节 马尔科夫决策模型简介·····	(361)
第六节 多目标决策·····	(363)

附表

附表 1	正态分布的面积和纵坐标	(370)
附表 2	χ^2 -分布的临界值表	(372)
附表 3	t-分布的临界值表	(373)
附表 4	F-分布的临界值表	(374)
附表 5	$F_{\max}$ 的临界值(哈特方差齐性检验)	(376)
附表 6	q-分布的临界值	(377)
附表 7	秩和检验表	(378)
附表 8	符号检验表	(379)
附表 9	H 检验表	(380)
附表 10	弗里德曼双向秩次方差分析 χ^2 值表	(382)
附表 11	泊松分布表	(384)
附表 12	随机数字表	(386)
附表 13	常用正交表	(387)
1. $L_8(2^7)$;	2. $L_8(4 \times 2^4)$;	3. $L_{16}(2^{15})$;
4. $L_{16}(4 \times 2^{12})$;	5. $L_{16}(4^2 \times 2^9)$;	6. $L_9(3^4)$;
7. $L_{18}(2 \times 3^7)$;	8. $L_{27}(3^{13})$;	9. $L_{25}(5^6)$.
参考书目		(393)

第一章 总 论

第一节 统计学的研究对象

随着科技进步和社会发展，统计方法作为定量分析的一种重要工具，越来越成为认识客观世界的重要手段而被广泛应用。

实用统计学既包括统计资料的收集、整理以及用一些概括的数字加以表示的统计方法，也包括介绍如何抽样、如何从样本出发对客观事物的未知部分进行推断和预测，直至为采取一定的决策提供有效的方法。

一、统计学的形成

统计学是阐明搜集、整理、分析大量现象的数字资料 and 进行推断的方法论科学。

统计学是在统计实践的基础上形成和发展起来的。它来源于统计学中的三个学派，即德国的国家记述学派，英国的政治算术学派和后期的数理统计学派，但主要是后两个学派。

（一）关于三个学派

国家记述学派的特征是对各个国家情况的比较叙述。它详细记载公共行政、司法、科学艺术等。在德国其主要人物之一是阿申瓦尔（Gatfried Achenwall 1719 年～1772 年）教授。他于 1746 年开始在德国马布里大学讲授《近代欧洲各国国势学概论》，以文字叙述的方法详述了英国、法国、德国、俄国、丹麦、瑞典、西

班牙、葡萄牙等国的物产、人口、土地、兵力等状况，并加以比较。他在讲义中根据拉丁文 status 创造了一个新名词——国势学或统计学（德语原文为 staatenkunde）。英文 statistics（统计学）是后来仿照 staatenkunde 构造的，意为研究政情的学问。由于这种统计学以文字叙述为特点，因而实在是“仅有统计学之名，而无统计学之实。”

英国威廉·配第（William Potty，1623 年～1687 年）所著《政治算术》和《爱尔兰的政治解剖》两部书，及其朋友格朗特（John Graunt，1620 年～1674 年）的《关于死亡表的自然和政治的观察》一书是政治算术学派的代表作。格朗特这本书对当时伦敦市的人口出生率、死亡率及性别比例进行计算，并利用数量对比关系推断其人口发展趋势。《政治算术》一书用大量数据对英、法、荷三国的经济实力进行了对比。尽管格朗特与配第都没有使用统计学一词，但这种以数量分析代替笼统叙述为特征的研究社会经济现象的理论与方法，日益受到重视，被看作是真正的统计学，并逐渐发展为现在的描述统计学。

继政治算术学派之后，兴起了数理统计学派。这一学派把概率论的思想方法引入统计学（政治算术意义上的统计学），用以研究自然和社会现象。在这方面工作最早的是法国数学家拉普拉斯（Laplace，1749 年～1827 年）。他在 1781 年的一篇论文里已开始用概率论分析男孩出生比例。作为学派创始人当推比利时学者阿道夫·凯特勒（Adolphe Quetelet，1796 年～1874 年）。他的主要著作有《论人和人的能力之发展，或社会物理学的经验》（1835 年）、《统计学的研究》（1844 年）、《关于概率论的书信》（1846 年）、《社会物理学》（1869 年）等。把概率论和统计学挂上钩，提出统计学的新的概念，开创了统计学新的发展领域。他肯定了表面上杂乱无章、偶然性占统治地位的社会现象如同自然现象一样，也有其内在的必然性。马克思对此给予了很高的评价。马克思说：“他曾立过大的功劳，论证即使社会生活表面的偶然事物，在它们

周期性的反复和它们周期性的平均中，也还具有内在的必然性”（马克思：《致库格曼书信集》，人民出版社1957年版，第79页）。

（二）关于概率论与数理统计学

概率的初步问题是从赌博中提出来的。17世纪中叶法国数学家帕斯卡（Pascal，1623年～1662年）和费马（Fermat，1601年～1665年）对此加以研究，为概率论的产生初步奠定了基础。18世纪初瑞士数学家贝努利（Bernoulli，1654年～1705年）又把它理论系统化地加以发展。直至1933年苏联数学家柯尔莫哥洛夫（А·И·Колмогоров）建立概率的公理化定义之后，概率论才真正成为近代数学的一个分支。

数理统计学的产生是19世纪中叶的事。从凯特勒把概率论和统计学挂上钩以后，引起很多人的兴趣，继续进行研究。有的人从数学上加以研究，有的人从自然科学中观察和对实验材料分析研究的应用上加以研究。凯特勒本人把概率论广泛应用于天文、物理、人口及人类社会的研究，并取得了丰富的成果。这个学科起初没有固定名称。1867年威梯斯坦发表了一篇关于数理统计的论文，从此定名为数理统计学。随后英国生物学家高尔登（Sir Francis Galton）于1889年提出回归与相关的概念，并发明相关系数，提出了矩法估计，建立了皮尔逊分布函数族及 χ^2 -分布，并发明了复相关和净相关的算法。稍后，奈曼（J·Neyman）和伊根·皮尔逊（Egon S·Pearson）建立了假设检验理论与区间估计理论。到20世纪20年代推断统计学的大样本理论趋于成熟。1908年戈赛特（William seely Gosset）以“学生”（Student）为笔名发表了一篇重要文章，提出了 t -分布，成为小样本理论研究的开端。数理统计作为一个数学学科的奠基者是英国学者费歇尔（R·A·Fisher），他提出 Z -分布，后经别人改进成为 F -分布，从而使小样本理论在数理统计中占据重要地位。另外他建立的“试验设计”也是科学家们普遍称赞的。20世纪30年代以来，推断统计学又出现一些新分支。如美国沃尔德（A·Wald）创立的序贯分析和决策

函数论；英国威沙特 (J · Wishart) 和美国霍太林 (Hotteling) 等发展了样本分布理论，奠定了多元统计的基础；柯克伦 (W · Cochran) 和考克斯 (C · M · Cox) 发展了抽样技术；休哈特 (W · A · Shewhart) 建立了产品质量控制；鲍克思 (George E · P · Box) 和詹金斯 (Gwilym M · Jenkins) 创建了现代时间序列分析；斯米尔诺夫 (H · B · Смирнов) 和柯尔莫哥洛夫等发展了非参数统计。

概率论和数理统计学都是研究随机现象数量规律性的数学学科。前者的任务是，从对随机现象的大量观察中提出随机现象的数学模型（分布），然后对这些分布进行研究，找出其内在的性质及其相互的联系；而数理统计学的任务却是研究怎样用有效的方法去收集和使用带有随机因素影响的数据，即研究如何抽样、如何从样本出发对总体的未知部分进行推断和预测，直至为采取一定的决策或行动提供依据。

随机现象大量存在，无论在社会界还是自然界。这是数理统计得到广泛应用的基础。近几十年来数理统计的广泛应用引人注目，尤其是在经济计量学方面。在经济关系中存在着许多相关的变量，彼此以某种数学模型联系着。我们的任务是决定这些模型中的参数。由于只能从样本中获得参数的估计值，就必须考虑误差及因素的作用，所以非采用随机模型不可，数理统计也就因此成为最基本的工具。

数理统计在任何一个领域里的应用，要想取得成功，都必须和那里的具体实际相结合。因为大数法则只是规律的表现形式而不是规律本身。规律本身只有从事物本身以及一事物与他事物的联系中去寻找。例如你若研究的是社会经济问题，就要以政治经济学为基础进行分析，从事物内部的相互联系中寻找其运动规律的本身。

二、统计学的研究对象

统计学的研究对象是大量现象的数量方面。但是，统计学所研究的不是事物的本质和规律本身，而是寻找事物的本质和规律的方法。正因为如此，统计学又称为统计方法。或者说，统计学是从数量方面研究现象的本质和规律的方法论科学。

(一) 统计学的主要内容

在利用统计学研究具体问题时，由于研究问题的目的和角度不同，所使用的统计方法也不同。但是，任何统计研究不外乎如下内容：

(1) 统计设计 主要解决抽样方法问题，即如何选取样本。

(2) 统计调查 选用一定的方法取得所需要的数据，即如何取到样本值。

(3) 统计整理 对调查所得到的数据进行系统化加工。

(4) 统计推断和预测 利用某些方法对事物总体作出估计或决策，或者对事物的未来发展进行预测。

(5) 统计解释 对推断或预测的合理性和可靠性加以说明。

(二) 统计的性质

统计是人类所进行的一种调查研究活动，是对一定的总体现象所进行的一种定量认识活动。与其他科学相比，它有如下特点：

(1) 数量性 任何事物都有质和量两个方面。统计学是从量的方面研究事物的，包括事物的数量特征、质量互变的数量界限及事物或现象间的数量关系。这一性质决定了它不同于哲学。虽然它们同属于认识事物的方法论科学，但哲学是关于如何从质的方面认识事物的科学，而统计学则是关于如何从量的方面认识事物的科学。

(2) 具体性 统计学所研究的是一定质下的量，而不是抽象的“纯数量”，所以它不同于纯数学。但是这并不妨碍数学运算和数学方法的应用。

(3) 大量性 统计学是通过大量观察、获取足够多的统计资料，以抵消偶然因素的影响，透过表象揭示事物的本质和规律的。没有大量数据的综合研究，就不是统计研究。由此可见，社会学中的方案研究和典型调查，显然不属于统计学。

(4) 广泛性 作为一门方法论科学，统计学的原理和方法适用于自然界和社会界，它是人类认识自然、认识社会的一种武器。局限于一定范围的应用统计学及其相关学科如经济计量学、人口统计学、社会统计学、生物统计学、管理统计学……，是统计学在应用领域里的自然延伸，但又有各自的特点。例如社会经济统计学，它除了具有统计学的一般特点外，还具有自身的特殊性，即它所研究的总体是社会经济现象，且大多数具有有限总体；它所研究的社会现象具有一定的阶级性。

第二节 统计学的学科体系及研究方法

统计学发展到今天已形成了庞大的体系，其研究方法独特且日趋完美。

一、统计学的学科体系

统计学分一般统计学、边缘统计学及统计学史三大分支。

在一般统计学中，由于研究问题的思想方法不同，又可分为描述统计学与推断统计学。

(一) 描述统计学

描述统计学是研究如何收集、整理和显示统计资料的方法的学科，它的特点是把研究对象作为确定性现象加以“描述”的。描述统计学又有描述统计学原理和应用描述统计学之分，前者是它的基础理论，包括统计设计理论、统计调查与整理理论、统计图表理论和统计指数理论等；后者包括人口统计、社会统计、生态统计、科技统计等。

（二）推断统计学

推断统计学又称数理统计学。它是把研究对象作为随机现象来处理的。它研究如何通过抽样搜集带有随机因素影响的数据，如何依此部分数据获取关于整体的信息，如何依此信息对整体进行推断或预测。推断统计学分理论推断统计学和应用推断统计学。前者包括抽样理论、估计与假设检验理论、相关与回归分析、多元统计分析 & 非参数统计理论等；后者主要有生物统计、天文统计、工程统计、医药统计及管理统计学等。

社会经济统计学中的工业统计、农业统计、商业统计、教育统计等，是推断统计与描述统计的结合应用。

边缘统计学有统计预测、对策、决策、仿真等。

二、统计学的研究方法

这里所指的方法不是该学科中的一个具体方法，而是从具体方法中概括出来的一般性方法或原则。它包括：

（一）大量观察法

由于受偶然因素的影响，个别现象之间往往存在着不同程度的数量差异，所以根据对个别现象的了解难以看出事物的本质或规律。如果进行大量观察，个别现象之间差异可以互相抵消，从而显示出事物的规律性来。例如考察某一城市的居民用电量，就个别住户而言，有的用多，有的用少，看不出规律；但就整个城市而言，每个月居民用电量几乎是一个常数。这说明统计学为了达到透过现象看本质的目的，必须用大量观察法，即对足够多的个别现象进行调查研究。这一方法是偶然性与必然性、个性与共性辩证统一在统计学中的成功运用。

大量观察法中“大量”以什么为限？应以“足够”为限。什么是足够？那是就研究的问题及采用的方法需要而言的，没有统一标准。它可以是所研究的全体个别现象，也可以是其中一部分。

（二）综合法、分析法

综合法是对大量观察所获资料加以“综合”，以反映事物的本质特征。常用的综合法有汇总法、平均法等。这种方法是对事物的描述。

分析法是对数据进行分解、对比，以研究事物的数量差异和数量关系的方法。分组分析法、因素分析法、相关与回归分析法、时间序列分析法等，都是统计学中常用的分析法。

（三）归纳推断法

归纳推断法是逻辑推断法的一种。它是由特殊到一般的推理方法。在统计学中，它是根据对已经获得的一定数量的个别现象的认识去推测或认识总体现象的本质或规律性的方法。归纳推断法是从开放系统中导出结论，其正确性不是必然的，而是基于一定概率意义下的。归纳推断法是现代统计学中的核心方法。

三、统计学的作用

统计学或统计方法是一种重要的定量分析方法。它不仅可以为理论研究提供工具，而且还可以为实际工作提供各种具体方法。其基本功能是：

（一）反映事物的数量状况

事物的数量状况包括数量特征、数量界限和数量关系等。要反映事物的数量状况，必须针对特定的问题，采用适当的方法。例如要了解一个乡的人口状况，乡政府采用描述统计的方法，调查了全乡各村的人口资料，计算出乡人口总数、各年龄段男女人口数量等。又如，国家要确定适度资产投资规模，需要研究投资率的数量界限。为此，可用描述统计方法根据经验资料求得一个合理的投资率；也可用推断统计学的回归分析模型找出合理投资率。再如，某农科所为了从五种油菜籽中找出适合本地的优良品种，将它们在同一土质、同样的中耕管理等条件下分别播种，以观察它们的产量。类似这样的数量关系，一般可用推断统计方法加以研究。

(二) 对事物的数量表现进行比较分析

统计学的第一个功能就是从原始资料中提取所需要的信息；其第二个功能就是对获得的信息作进一步的分析比较。例如进一步比较全乡的男女比例、计算抚养比（0岁至14岁人口数与65岁及65岁以上人口数之和除以15岁至64岁人口数），和少年儿童系数（14岁以下人口数与总人口数之比），并和其他乡进行比较。通过比较找出差异产生的原因。

(三) 根据事物数量资料作出推断

统计学的重要任务之一是根据部分资料对整体进行推断。例如，根据五种油菜籽的产量试验结果，能否断定它们的平均产量有无显著差异。

(四) 利用事物的数量关系对未来进行预测

例如利用时间序列模型由今天的资料可以推测未来的状况。或者由回归模型中的自变量预测因变量。例如由移算法，根据现在的人口资料可预测未来的人口状况。

以上四点功能是统计学本身固有的，具有相对稳定性。统计学在各具体领域里的应用，可以进一步显示出它的职能或作用。

统计学的作用不容置疑，但也决不是万能的。我们反对那种为达到特定目的而以个别事例为依据的欺骗行为。美国著名作家马克·吐温将世间撒谎者分成三等：撒谎者、该下地狱的撒谎者和统计学家（指滥用、误用统计方法的人）。

第三节 统计学的基本概念

一、总体、个体、样本

(一) 总体

研究对象的全体称为**总体**。由于对“研究对象”理解不同，可得出如下两个不同的定义：以描述统计观点看，总体是所研究的