



现代经济与管理类规划教材

统计学 教程

卢小广 编著



清华大学出版社 · 北京交通大学出版社

◇现代经济与管理类规划教材

统计学教程

卢小广 编 著

清华大学出版社
北京交通大学出版社

· 北京 ·

内 容 简 介

按照经济管理学科专业的要求,本书从描述统计和推断统计两个层次上,系统地综述了统计学的基本理论、基本原理和基本方法。在严谨科学的理论框架下,本书强调简洁性、系统性和可读性,力求深入浅出,理论联系实际,着重提高读者运用各种统计方法和统计软件分析问题、解决问题的实际能力,为经济管理专业课程的学习和研究奠定基础。

本书为各类经济管理学科的研究生、本科生教学用书,也可以作为从事经济管理的研究人员和工作人员的专业参考书籍。

版权所有,翻印必究。举报电话:010-62782989 13501256678 13801310933

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

本书防伪标签采用特殊防伪技术,用户可通过在图案表面涂抹清水,图案消失,水干后图案复现;或将表面膜揭下,放在白纸上用彩笔涂抹,图案在白纸上再现的方法识别真伪。

图书在版编目(CIP)数据

统计学教程/卢小广编著. — 北京:清华大学出版社;北京交通大学出版社,2005.12
(现代经济与管理类规划教材)

ISBN 7-81082-640-9

I. 统… II. 卢… III. 统计学-高等学校-教材 IV. C8

中国版本图书馆CIP数据核字(2005)第118253号

责任编辑:吴嫦娥

出 版 者:清华大学出版社 邮编:100084 电话:010-62776969 <http://www.tup.com.cn>

北京交通大学出版社 邮编:100044 电话:010-51686414 <http://press.bjtu.edu.cn>

印 刷 者:北京瑞达方舟印务有限公司

发 行 者:新华书店总店北京发行所

开 本:185×230 印张:22 字数:493千字

版 次:2006年1月第1版 2006年1月第1次印刷

书 号:ISBN 7-81082-640-9/C·12

印 数:1~4 000册 定价:29.00元

本书如有质量问题,请向北京交通大学出版社质监组反映。对您的意见和批评,我们表示欢迎和感谢。

投诉电话:010-51686043, 51686008; 传真:010-62225406; E-mail: press@center.bjtu.edu.cn。

前言

面对今天的知识经济时代,信息呈几何级数迅猛增长,各类信息的采集、处理、辨识和分析已经成为当今各行各业日常工作的重要内容,作为数据处理和分析技术的统计方法越来越广泛地应用于科学研究、生产管理、经济分析和经营决策的各个方面。

统计学是国家教育部规定的财经类专业的核心课程,在我国高等院校的经济学科和管理学科中普遍开设了这一课程。本书综合论述了统计学的基本理论和基本方法。基于经济管理学科专业基础课程的考虑,着重介绍了经济管理方面所涉及的统计学理论和方法。本书可供经济管理学科教学使用,也可以作为从事经济管理工作人员的参考书籍。

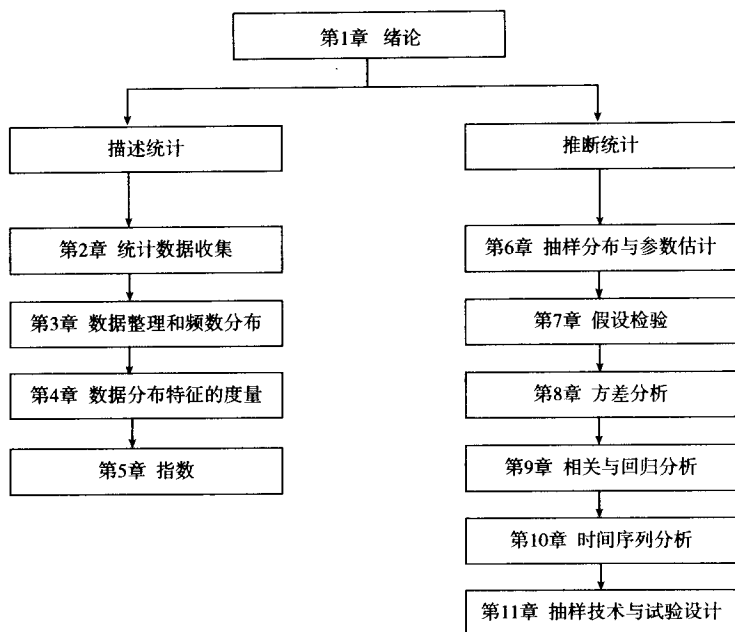
本书具有以下特点。

(1) 体系结构上贯彻“大统计”的观点。根据“统计学是研究如何搜集数据,分析数据,以便得出正确认识结论的方法论科学”,统计学是一门既适用于自然科学,又适用于社会科学的通用的方法论科学的原则,从统计数据出发,以统计数据的处理和分析为核心,在描述统计和推断统计两个层次上,论述统计学的基本理论和基本方法。

(2) 内容布局上注重统计方法的严谨性和实用性。在确保理论体系科学严谨的前提下,力求体现本书的简洁性、系统性和可读性特征,以突出统计方法的实用性本质,来适应学生 and 实际工作者自学的需要。例如,本书删去了概率论的章节,因为概率论在各高校已经开设成独立的课程,没有必要在统计学中重复讲述。又如,本书在论述样本分布的章节中始终强调独立同分布样本概念,并将融合了样本分布、参数估计、假设检验等知识的抽样技术,安排在最后的“抽样技术与实验设计”一章中介绍,以避免出现概念混乱。

(3) 方法掌握上强调计算机软件的使用。作为非统计专业的读者,没有必要掌握全部统计方法的推导和论证过程,关键在于学会运用各种统计方法分析问题,研究问题,尤其是正确地运用各种统计方法,采用各种数据处理和统计分析软件来进行统计分析的实际能力。所以本书专门设立了 Excel 应用栏目,采用目前使用最普遍的 Microsoft Excel 电子表格软件,结合案例讲解具体的统计方法,重点介绍了 Excel 的“分析工具库”和有关统计函数等内容。

本书的章节结构如以下框图所示。



描述统计和推断统计是本书对统计学知识的基本分类，但在具体章节安排中也有少量的交叉。例如，在第2章的“2.2.3 敏感问题的随机化处理”属于推断统计的内容，以及第10章的“10.1 描述性分析”属于描述统计的范畴。

在本书的编写过程中，参考了国内外有关学者的专著和论文；并得到了北京交通大学出版社吴嫦娥编辑的支持，在此一并致谢。

本书是编者在统计学教学和研究过程之中，长期思考不断探索的一个成果，也是力求完善统计学教材的一次尝试。种种不当和疏漏在所难免，恳请各位同行和读者提出宝贵意见，以便再版时修正。联系方式：hongrui@jlonline.com

编 者
2006. 1

目 录

第1章 绪论	(1)
1.1 统计与统计学	(1)
1.1.1 统计实践的出现	(1)
1.1.2 统计学的产生和发展	(1)
1.1.3 统计学的定义	(3)
1.1.4 统计一词的涵义	(4)
1.1.5 描述统计和推断统计	(4)
1.2 统计的几个基本概念	(4)
1.2.1 总体和样本	(4)
1.2.2 参数和统计量	(6)
1.2.3 数据计量尺度	(6)
1.2.4 变量	(7)
1.2.5 统计指标和标志	(8)
◇ Excel 应用	(10)
◇ 本章小结	(11)
◇ 思考与练习	(12)
◇ 人物小传	(12)
第2章 统计数据收集	(13)
2.1 统计数据的来源	(13)
2.1.1 统计调查	(13)
2.1.2 数据的搜集方法	(16)
2.2 统计调查设计	(17)
2.2.1 统计调查方案设计	(17)
2.2.2 问卷设计	(18)
2.2.3 敏感问题的随机化处理	(22)
2.3 统计数据质量	(23)
2.3.1 统计数据的误差种类	(23)
2.3.2 统计数据的质量标准	(24)

◇ Excel 应用	(24)
◇ 本章小结	(27)
◇ 思考与练习	(27)
◇ 人物小传	(28)
第3章 数据整理和频数分布	(29)
3.1 数据的审核和筛选	(29)
3.1.1 数据的审核	(29)
3.1.2 数据的筛选	(30)
3.1.3 数据的排序	(31)
3.2 数据的分组	(31)
3.2.1 数据分组的意义	(31)
3.2.2 非数值型数据的分组	(31)
3.2.3 数据型数据的分组	(33)
3.2.4 组中值	(36)
3.3 数据的频数分布	(38)
3.3.1 频数与频数分布	(38)
3.3.2 累积频数	(38)
3.3.3 异距分组与标准组距频数	(39)
3.4 绝对数	(39)
3.4.1 绝对数的概念	(40)
3.4.2 绝对数的种类	(40)
3.4.3 绝对数的计量单位	(41)
3.5 数据的展示	(42)
3.5.1 统计表	(42)
3.5.2 统计图	(44)
◇ Excel 应用	(46)
◇ 本章小结	(54)
◇ 思考与练习	(54)
◇ 人物小传	(55)
第4章 数据分布特征的度量	(56)
4.1 集中趋势的度量	(56)
4.1.1 众数	(56)
4.1.2 中位数	(58)

4.1.3	均值	(61)
4.1.4	调和平均数	(65)
4.1.5	几何平均数	(67)
4.1.6	均值、调和平均数和几何平均数的比较	(68)
4.1.7	均值、众数和中位数的比较	(69)
4.2	离散程度的度量	(71)
4.2.1	异众比率	(71)
4.2.2	四分位差	(71)
4.2.3	取值范围	(72)
4.2.4	平均差	(73)
4.2.5	方差和标准差	(75)
4.2.6	离散系数	(78)
4.2.7	标准化值	(79)
4.3	偏态和峰度的度量	(81)
4.3.1	偏态系数	(81)
4.3.2	峰度系数	(82)
◇	Excel 应用	(84)
◇	本章小结	(93)
◇	思考与练习	(94)
◇	人物小传	(95)
第5章	指数	(96)
5.1	相对数	(96)
5.1.1	相对数的量纲	(97)
5.1.2	计划完成相对数	(98)
5.1.3	结构相对数	(99)
5.1.4	比例相对数	(100)
5.1.5	比较相对数	(100)
5.1.6	强度相对数	(101)
5.1.7	动态相对数	(101)
5.1.8	相对数小结	(102)
5.2	综合指数	(103)
5.2.1	指数的分类	(103)
5.2.2	综合指数原理	(104)
5.2.3	拉氏指数和帕氏指数	(106)

5.2.4 著名的指数公式	(108)
5.2.5 拉氏公式和帕氏公式的变形	(109)
5.3 平均指数	(110)
5.3.1 平均指数数量指数和质量指数	(110)
5.3.2 固定权数的平均指数	(111)
5.4 指数体系	(113)
5.4.1 指数体系与因素分析	(113)
5.4.2 指数体系的动态分析	(114)
5.4.3 指数体系的静态分析	(115)
5.5 我国的物价指数	(118)
5.5.1 居民消费价格指数	(118)
5.5.2 商品零售价格指数	(120)
5.5.3 工业品出厂价格指数	(121)
5.5.4 股票价格指数	(123)
◇ Excel 应用	(123)
◇ 本章小结	(125)
◇ 思考与练习	(125)
◇ 人物小传	(127)

第6章 抽样分布与参数估计	(128)
6.1 抽样分布	(128)
6.1.1 总体、个体和样本	(128)
6.1.2 大数定律和中心极限定理	(129)
6.1.3 三种分布	(130)
6.1.4 样本均值的抽样分布	(131)
6.1.5 样本比例的抽样分布	(134)
6.1.6 样本方差的抽样分布	(135)
6.1.7 两个总体样本统计量的抽样分布	(136)
6.2 参数估计的一般问题	(138)
6.2.1 估计量和估计值	(138)
6.2.2 点估计	(138)
6.2.3 点估计量的评价准则	(140)
6.2.4 区间估计	(142)
6.3 单一总体参数的区间估计	(144)
6.3.1 总体均值的区间估计	(144)

6.3.2 总体比例的区间估计	(149)
6.3.3 总体方差的区间估计	(150)
6.4 两个总体参数的区间估计	(150)
6.4.1 两个总体均值之差的区间估计	(150)
6.4.2 两个总体比例之差的区间估计	(154)
6.4.3 两个总体方差比值的区间估计	(155)
◇ Excel 应用	(156)
◇ 本章小结	(158)
◇ 思考与练习	(159)
◇ 人物小传	(161)
第7章 假设检验	(162)
7.1 假设检验的一般问题	(162)
7.1.1 小概率原理	(162)
7.1.2 假设检验的一般步骤	(163)
7.1.3 假设检验两类错误	(164)
7.1.4 双侧检验和单侧检验	(166)
7.1.5 p 值	(168)
7.2 单一总体参数的假设检验	(169)
7.2.1 总体均值的假设检验	(169)
7.2.2 总体比例的假设检验	(174)
7.2.3 总体方差的假设检验	(175)
7.3 两个总体参数的假设检验	(176)
7.3.1 两个总体均值之差的假设检验	(176)
7.3.2 两个总体比值之差的假设检验	(180)
7.3.3 两个总体方差之比的假设检验	(180)
◇ Excel 应用	(182)
◇ 本章小结	(186)
◇ 思考与练习	(188)
◇ 人物小传	(189)
第8章 方差分析	(190)
8.1 方差分析的一般问题	(190)
8.1.1 方差分析的基本原理	(190)
8.1.2 方差分析的基本假定	(192)

8.2 单因素方差分析	(192)
8.2.1 单因素方差分析的步骤	(192)
8.2.2 方差分析表	(196)
8.2.3 方差分析中的多重比较	(197)
8.3 双因素方差分析	(199)
8.3.1 无交互作用的双因素方差分析	(199)
8.3.2 有交互作用的双因素方差分析	(203)
◇ Excel 应用	(206)
◇ 本章小结	(213)
◇ 思考与练习	(213)
◇ 人物小传	(216)
第9章 相关与回归分析	(217)
9.1 相关关系	(217)
9.1.1 相关关系的概念	(217)
9.1.2 相关关系的度量	(219)
9.2 一元线性回归	(222)
9.2.1 一元线性回归模型	(222)
9.2.2 一元线性回归方程的最小二乘估计	(223)
9.2.3 一元线性回归方程的拟合优度	(225)
9.2.4 一元线性回归方程的显著性检验	(227)
9.2.5 运用回归方程进行估计	(229)
9.3 多元线性回归	(233)
9.3.1 多元线性回归模型	(233)
9.3.2 多元线性回归的最小二乘估计	(234)
9.3.3 多元线性回归方程的拟合优度	(236)
9.3.4 多元线性回归方程的显著性检验	(238)
9.3.5 运用多元线性回归方程进行估计	(240)
9.4 非线性回归的线性化	(243)
◇ Excel 应用	(245)
◇ 本章小结	(252)
◇ 思考与练习	(252)
◇ 人物小传	(254)

第 10 章 时间序列分析	(255)
10.1 描述性分析	(255)
10.1.1 时间序列的种类	(255)
10.1.2 发展水平和平均发展水平	(256)
10.1.3 增长量与平均增长量	(259)
10.1.4 发展速度和增长速度	(260)
10.1.5 平均发展速度与平均增长速度	(261)
10.2 长期趋势分析	(263)
10.2.1 长期趋势的因素分析	(263)
10.2.2 移动平均法	(264)
10.2.3 指数平滑	(267)
10.2.4 模型拟合法	(270)
10.3 季节变动分析	(285)
10.3.1 长期趋势的剔除	(285)
10.3.2 季节指数的计算	(287)
10.4 循环变动分析	(289)
◇ Excel 应用	(291)
◇ 本章小结	(295)
◇ 思考与练习	(295)
◇ 人物小传	(297)
第 11 章 抽样技术与试验设计	(298)
11.1 抽样技术	(298)
11.1.1 抽样技术的几个基本概念	(298)
11.1.2 简单随机抽样	(299)
11.1.3 系统抽样	(301)
11.1.4 分层抽样	(303)
11.1.5 整群抽样	(307)
11.1.6 必要样本容量	(308)
11.1.7 根据费用函数确定样本容量	(311)
11.2 试验设计	(313)
11.2.1 一种简单的试验设计	(313)
11.2.2 随机化区组设计	(314)
11.2.3 拉丁方设计	(316)
◇ 本章小结	(319)

◇ 思考与练习	(319)
◇ 人物小传	(322)
附录 A 常用统计表	(323)
表 A.1 标准正态分布表	(323)
表 A.2 正态分布常用分位数表	(325)
表 A.3 t 分布表	(325)
表 A.4 χ^2 分布表	(326)
表 A.5 F 分布表	(328)
习题参考答案	(330)
参考文献	(338)

第 1 章

绪 论

在工作、生活、学习中,人们经常接触各类统计数据、统计图表,需要搜集、处理、分析和使用各类统计数据,统计已经渗入人类社会经济生活的每个角落。

本章主要介绍统计学的产生和发展,以及统计学的一些基本概念。

1.1

统计与统计学

1.1.1 统计实践的出现

随着人类文明的出现,远在原始社会初期的氏族、部落在安排狩猎、分配食物时,就有了最初的计数活动,孕育着统计实践的萌芽。

统计实践活动随着国家的产生而不断加强和完善。早在公元前 3 000 多年,古埃及为了建造金字塔征集经费、劳力,对全国的财产、人口进行了数次普查。我国在春秋战国时期,管仲就明确指出“不明于计数,犹如无舟楫欲径于水,险也”;商鞅提出了“强国知十三数”,系统地列出了治理国家必须了解的十三项指标,包括人口、土地、粮食等。

统计是人类的一项社会实践活动。不论何时何地,只要提到统计,人们总是要把它与数据的总计、平均、差异比较,和大量数据的综合分析联系起来。统计是通过对事物的数量特征进行搜集、整理和分析,进而得出对事物规律性认识的社会实践活动。各类事物及其现象的发展变化规律,都表现为质与量的辩证统一,要认识客观存在的事物及其现象的规律性,就必须把握其数量关系的基本特征及度的界限,这些都离不开统计。

1.1.2 统计学的产生和发展

在数千年的统计实践活动中,人们对统计规律的认识逐渐加深,通过对统计实践活动的

长期总结和逐步归纳,将其系统化为一门规范的科学体系——统计学。统计学作为一门独立的科学已有三百多年的历史,一般将统计学的产生和发展分为三个阶段。

1. 古典统计学时期(17世纪中叶至18世纪末)

17世纪中叶欧洲各国相继进入资本主义经济发展阶段,处于思想活跃的社会变革时期。各国不约而同地从不同领域开始了统计学的奠基工作,并相继形成了统计学的三大来源,即政治算术学派、国势学派和古典概率理论,为近代统计学奠定了理论基础,这一阶段是统计学的萌芽时期。

政治算术学派产生于17世纪中叶的英国,代表人物是威廉·配第(William Petty)和约翰·格朗特(Johan Graunt)。英国古典政治经济学的创始人,威廉·配第博士在1690年出版了《政治算术》一书,采用“数字、重量、尺度”等定量分析工具,用数量比较的方法,进行经济政治研究。约翰·格朗特在1662年发表了《关于死亡表的自然观察与政治观察》一书,建立了最初的人口增长模型,对伦敦市的人口现象进行了统计研究,提出了“大量恒静”这一著名的格言。

国势学派产生于18世纪的德国。1660年康令(Hermann Conring)博士把国势学从法学、史学、地理学等学科中独立出来,在大学中讲授“实际政治家所必需的知识”。后由休姆采尔(Martin Schneitzel)将其更名为《政治学·统计学讲义》。阿痕瓦尔(Gottfried Achenwall)于1749年正式确定了统计学这一学科的名称及有关统计学的一些术语。

古典概率论的研究始于16世纪的意大利,直到17世纪中叶才得到一般化的解法。并在18世纪的法国、瑞士等国得到广泛发展,最终于19世纪初叶由法国数学家、统计学家拉普拉斯(P. S. Laplace, 1749—1827)在总结前人成果的基础上,出版了名著《概率论分析理论》一书,从而形成完整的应用理论体系。

2. 近代统计学时期(19世纪初至20世纪初)

近代统计学的主要贡献是建设和完善了统计学的理论体系,并逐渐形成了以随机现象的推断统计为主要内容的数理统计学和以传统的政治经济现象的描述统计为主要内容的社会统计学两大学派。

数理统计学派的代表人物比利时的凯特勒(Lambert Adolphe Jacques Quetelet)博士深受拉普拉斯影响,在其《社会物理学》中将概率论引入统计学,在生物统计和天文学研究中成功地运用了二项分布和正态分布方法。凯特勒博士认为要促进科学的发展,就必须更多地应用统计学,大力提倡采用统计方法对自然现象和社会现象的规律性进行观察。

社会统计学派的代表人物有克尼斯(K. G. A. Knies),梅尔(Georg Von Mayr)和恩格尔(Christian Lonrenz Ernst Engel)等。社会统计学派继承和发展了“政治算术学派”的理论体系,完善了统计有关描述性的分析方法,建立了较为完整的统计指标方法和社会统计学科体系,对当时的政府统计工作和统计学理论的推广产生了积极的促进作用。

3. 现代统计学的发展时期（20 世纪初到现在）

自 20 世纪初到 50 年代是数理统计学派迅速发展的时期。由费歇尔、卡尔·皮尔逊、奈曼、瓦尔德等为代表的一代统计学大师，共同完成了现代统计学的理论框架，使统计学成为一门严谨的学科，同时数理统计学逐渐成为了主流统计学。

皮尔逊（Karl Pearson）定义了相关系数 ρ 的联合正态分布，估计误差的联合分布，以及关于拟合度的 χ^2 检验；费歇尔（Ronald Aylmer Fisher）确立了统计推测理论，样本分布理论，试验计划法及 F 分布理论，特别是于 1921 年发表了“理论统计学的数学基础”一文，一般性地规定了数理统计学的基本架构；1937 年奈曼（Jerzy Neyman）发表了关于置信区间的奠基性论文；1950 年瓦尔德（Wald）发表“统计决策函数”。至此，数理统计学已经形成了一个统一的理论框架，并一直沿用至今。

1950 年以后，受计算机和新兴科学的影响，这一时期统计学的研究和应用范围越来越广，如科克伦（W. G. Cochran）的实验设计和抽样理论、安得森（T. W. Anderson）的复变数分析等。

20 世纪中，统计学在社会经济领域里的应用首推 1953 年创立的 SNA（A System of National Accounts and Supporting Tables，国民账户体系及辅助表）核算体系，以及 1968 年发表的新 SNA 体系和 1993 年经过再次改进的 SNA 体系。在这一领域中，著名的统计学家理查德·斯通（Richard Stone）和瓦西里·列昂惕夫（Wassily Leontief）两位诺贝尔经济学奖得主做出了杰出贡献。

✦ 讨论题 结合统计学产生和发展的历史，提出你对统计学的认识。

1.1.3 统计学的定义

从统计实践和统计学的产生和发展可以看出，统计学是一种应用领域宽广，分支众多，学派林立的大学科。因此，简单地对统计学给出一个普遍赞同的定义是困难的。《大不列颠百科全书》给出的定义——统计学是“用以收集数据、分析数据和由数据得出结论的一组概念、原则和方法”。1998 年 7 月国家教育部“普通高等学校本科专业目录和专业设置”中给出的定义为“统计学是研究如何搜集数据，分析数据，以便得出正确认识结论的方法论科学”。

具体来说，可以从以下三个方面来把握统计学的定义。

首先，统计学是方法论科学，而不是实质性科学。统计学不研究具体规律本身，而是研究普遍存在的数量关系的计量及其分析，并通过数量分析来认识特定事物内在规律性的方法。

其次，统计学的应用范围不仅限于社会科学，也不只限于自然科学。统计学方法来自于社会科学和自然科学的统计实践活动，可以用于社会现象也可以用于自然现象，统计学是一

种通用的方法论科学。统计学是一门不依附于任何实质性科学而存在的方法论，它是一门独立的方法论科学。

第三，统计学的研究对象既包括确定性现象的总体数量特征，也包括随机现象的总体数量特征，即统计学是研究各类事物及其现象的数量特征及其内在联系的方法论科学。

1.1.4 统计一词的涵义

“统计”一词一般包括如下三个方面涵义。

一是统计工作，即对事物及其现象的数量特征进行数据搜集、整理和分析的社会实践活动过程。

二是统计数据，一般也称为统计资料，即统计工作的成果。

三是统计科学，即对统计工作规律进行总结和理论概括的系统学说。

统计学与统计工作之间是理论和实践的关系。统计学来源于统计实践的经验总结，同时作为系统化了的的人类知识积累，又指导着人们有效地开展统计实践，并在统计实践中得到检验和修正，推动统计学自身的完善和发展。统计数据是统计工作的成果，不仅反映了统计实践活动的规模 and 水平，同时还反映了统计科学的完善和规范，为检验和发展统计学理论和方法提供了重要依据和必要中介。

1.1.5 描述统计和推断统计

描述统计 (Descriptive Statistics) 是研究数据的收集、整理、综合及分析的统计方法体系。

推断统计 (Inferential Statistics) 是研究利用样本数据估计、检验及分析总体特征的统计方法体系。描述统计和推断统计构成完整的现代统计学方法体系。



1.2

统计的几个基本概念

1.2.1 总体和样本

总体 (Population) 是指由所研究的全部个体所构成的集合。构成总体的每一元素称为个体 (Unit)，或总体单位 (Population Unit)。

总体是统计最基本的要素和的概念。总体是统计研究的对象，统计认识活动的目的决定