

★高等院校精品教材

统计学

STATISTICS

罗 斐◎主编



知识产权出版社

全国百佳图书出版单位

☆高等院校精品教材

统 计 学

罗 斐 主编



知识产权出版社

全国百佳图书出版单位

内容提要

本书是在本人多年的讲稿基础上,以统计理论和实践相结合为依据,借鉴国内外相关教材和资料,适当增减有关内容,以适应学习型课堂建设需要而编写的。全书共分12章,前9章是基础也是重点,后3章是宏观方面和应用方面的补充。本书的主要内容包括:统计学概论,统计数据的描述,概率基础与抽样分布,参数估计,假设检验,方差分析与试验设计,相关分析与回归分析,时间序列分析,统计指数,国民经济统计,常用统计软件及其应用。本书可作为高等学校经济学、工商管理、市场营销、会计、国际贸易等专业的本科生教材,也可作为统计研究人员、统计从业人员或其他相关管理人员的参考书。

责任编辑:王辉

责任出版:刘译文

图书在版编目(CIP)数据

统计学/罗斐主编. —北京:知识产权出版社,2013.11

ISBN 978-7-5130-2409-9

I. ①统… II. ①罗… III. ①统计学—高等学校—教材 IV. ①C8

中国版本图书馆 CIP 数据核字(2013)第 267733 号

统计学

TONG JI XUE

罗斐 主编

出版发行:知识产权出版社

社 址:北京市海淀区马甸南村1号

网 址:<http://www.ipph.cn>

发行电话:010-82000893 82000860 转 8101

责编电话:010-82000860 转 8381

印 刷:北京富生印刷厂

开 本:787mm×1092mm 1/16

版 次:2013年12月第1版

字 数:480千字

邮 编:100088

责编传真:010-82000860 转 8353

传 真:010-82000893

责编邮箱:wanghui@cnipr.com

经 销:新华书店及相关销售网点

印 张:21.75

印 次:2014年1月第2次印刷

定 价:48.00元

ISBN 978-7-5130-2409-9

出版权专有 侵权必究

如有印装质量问题,本社负责调换。

前 言

《统计学》作为高等院校经济管理类专业的核心课程，具有很强的基础性和应用性。统计“要用数据说话”，揭示客观事物的数量规律性，为社会经济生活服务，因而重在应用。统计教学的总体目标：一是让学生系统地掌握统计学的理论和方法；二是培养学生理论联系实际，分析问题和解决问题的能力；三是尝试“学习型统计课堂建设”，培养学生的终身学习能力和创新能力。

本书具有以下特色：

(1) 覆盖面广。本书在结构、内容上，按照学科特点和统计研究的基本环节来安排，既突出重点又照顾一般，既重视理论又强调实践，既适合统计专业又适合非统计专业。

(2) 适用性强。本书既适用于经济学、工商管理类大学本专科生，又适用于人文社会科学及其他有关人士作为教材或参考书。

(3) 实践性强。本书在本人多年的讲稿和教学体会的基础上，参考国内外有关专家的文献资料，结合例题、习题和案例分析，强化学生的实践操作能力。

在本书的编写过程中，得到了我的研究生的大力支持，参考了国内外有关专家、学者的教材和著作，在此一并表示衷心的感谢！

由于编者水平有限，本书中仍有疏漏甚至错误之处，恳请同行和读者提出批评和建议，以利进一步修改和完善。

编者

2013年11月于北京

目 录

第 1 章 统计学概论	(1)
1.1 统计数据与统计学	(1)
1.2 统计学的产生和发展	(2)
1.3 统计学的研究对象	(3)
1.4 统计学的分科	(4)
1.4.1 描述统计和推断统计	(4)
1.4.2 理论统计和应用统计	(5)
1.5 统计学的基本概念	(6)
1.5.1 总体	(6)
1.5.2 总体单位	(6)
1.5.3 指标	(6)
1.5.4 标志	(7)
1.5.5 变量	(7)
1.5.6 样本	(7)
1.6 统计研究的基本环节	(7)
1.6.1 统计设计	(8)
1.6.2 统计调查	(8)
1.6.3 整理与分析	(9)
1.6.4 统计应用	(9)
本章小结	(10)
思考与练习	(10)
案例分析 20 世纪统计方法对质量改进的三大贡献	(10)
第 2 章 统计数据的描述	(13)
2.1 数据的计量尺度	(13)
2.1.1 列名尺度	(13)
2.1.2 顺序尺度	(13)
2.1.3 间隔尺度	(13)
2.1.4 比例尺度	(14)

2.2	统计数据的来源	(14)
2.2.1	间接获取的数据	(14)
2.2.2	直接获取的数据	(15)
2.3	统计数据的质量	(15)
2.4	统计数据的整理	(16)
2.4.1	统计数据的分组	(16)
2.4.2	次数分配	(17)
2.4.3	次数分配直方图	(19)
2.4.4	洛伦茨曲线与基尼系数	(21)
2.5	分布集中趋势的测度	(22)
2.5.1	众数	(22)
2.5.2	中位数	(23)
2.5.3	分位数	(24)
2.5.4	均值	(25)
2.5.5	几何平均数	(27)
2.5.6	切尾均值	(28)
2.5.7	众数、中位数和均值的关系	(29)
2.6	分布离散程度的测量	(31)
2.6.1	极差	(31)
2.6.2	内距	(31)
2.6.3	方差和标准差	(32)
2.6.4	离散系数	(33)
2.7	分布偏态与峰度的测度	(34)
2.7.1	偏态及其测度	(34)
2.7.2	峰度及其测度	(36)
2.8	统计表和统计图	(36)
2.8.1	统计表	(37)
2.8.2	统计图	(38)
	本章小结	(45)
	思考与练习	(46)
	案例分析 投资组合选择	(50)
第3章	概率基础与抽样分布	(52)
3.1	概率基础	(52)
3.1.1	随机现象与随机事件	(52)
3.1.2	概率的性质及其计算	(52)

3.1.3 随机变量及其分布	(55)
3.2 抽样分布	(60)
3.2.1 常用的抽样方法	(60)
3.2.2 简单随机抽样	(60)
3.2.3 分层抽样	(60)
3.2.4 系统抽样	(61)
3.2.5 整群抽样	(61)
3.2.6 抽样分布的概念	(62)
3.2.7 $\bar{x}$ 抽样分布的形式	(64)
3.2.8 $\bar{x}$ 抽样分布的特征	(66)
3.2.9 样本比率的抽样分布	(69)
3.2.10 样本方差的抽样分布	(70)
3.2.11 两个样本统计量的抽样分布	(71)
3.3 中心极限定理的应用	(73)
本章小结	(75)
思考与练习	(76)
案例分析 通过多样化减少投资风险	(78)
 第4章 参数估计	 (80)
4.1 参数估计的基本原理	(80)
4.1.1 估计量与估计值	(80)
4.1.2 点估计与区间估计	(80)
4.1.3 评价估计量的标准	(83)
4.2 一个总体参数的区间估计	(84)
4.2.1 总体均值的区间估计	(84)
4.2.2 总体比率的区间估计	(87)
4.2.3 总体方差的区间估计	(89)
4.3 两个总体参数的区间估计	(90)
4.3.1 两个总体均值之差的区间估计	(90)
4.3.2 两个总体比率之差的区间估计	(94)
4.3.3 两个总体方差比的区间估计	(95)
4.4 样本容量的确定	(98)
4.4.1 估计总体均值时样本容量的确定	(98)
4.4.2 估计总体比率时样本容量的确定	(99)
4.4.3 估计两个总体均值之差的样本容量的确定	(100)
4.4.4 估计两个总体比率之差时样本容量的确定	(100)

本章小结	(101)
思考与练习	(101)
案例分析 估计灯泡的使用寿命	(104)
第 5 章 假设检验	(106)
5.1 假设检验的基本问题	(106)
5.1.1 假设的陈述	(106)
5.1.2 两类错误与显著性水平	(107)
5.1.3 检验统计量与拒绝域	(108)
5.2 一个总体参数的检验	(110)
5.2.1 总体均值的检验	(110)
5.2.2 总体比率的检验	(114)
5.2.3 总体方差的检验	(115)
5.3 两个总体参数的检验	(117)
5.3.1 两个总体均值之差的检验	(117)
5.3.2 两个总体比率之差的检验	(123)
5.3.3 两个总体方差比的检验	(126)
本章小结	(127)
思考与练习	(128)
案例分析 新饲料养牛计划的有效性检验	(132)
第 6 章 方差分析与试验设计	(134)
6.1 方差分析引论	(134)
6.1.1 方差分析的内涵	(134)
6.1.2 方差分析的思路	(135)
6.1.3 方差分析的假定	(137)
6.1.4 问题的一般表述	(138)
6.2 单因素方差分析	(138)
6.2.1 数据结构	(138)
6.2.2 分析步骤	(139)
6.2.3 用 Excel 进行方差分析	(144)
6.2.4 关系强度的测量	(145)
6.3 方差分析中的多重比较	(146)
6.3.1 多重比较的目的	(146)
6.3.2 多重比较的方法	(146)
6.4 双因素方差分析	(148)

6.4.1 双因素方差分析及其类型	(148)
6.4.2 无交互作用的双因素方差分析	(149)
6.4.3 有交互作用的双因素方差分析	(154)
6.5 试验设计	(159)
6.5.1 完全随机化设计	(159)
6.5.2 随机化区组设计	(161)
6.5.3 因子设计	(162)
本章小结	(163)
思考与练习	(164)
案例分析 检验募捐方式的有效性	(166)
 第7章 相关分析与回归分析	(168)
7.1 相关与回归的基本概念	(168)
7.1.1 变量间的相互关系	(168)
7.1.2 相关关系的类型	(168)
7.1.3 相关分析与回归分析	(170)
7.2 简单线性相关与回归分析	(170)
7.2.1 简单线性相关系数及检验	(170)
7.2.2 回归函数	(173)
7.2.3 回归系数的估计	(175)
7.2.4 简单线性回归模型的检验	(179)
7.2.5 一元线性回归模型的预测	(182)
7.3 多元线性回归分析	(184)
7.3.1 多元线性回归模型及假设	(184)
7.3.2 多元线性回归模型的估计	(185)
7.3.3 多元线性回归模型的检验	(188)
7.4 非线性相关与回归分析	(191)
7.4.1 非线性回归的函数形式与估计方法	(192)
7.4.2 非线性相关指数	(197)
本章小结	(197)
思考与练习	(198)
案例分析 如何控制不良贷款	(202)
 第8章 时间序列分析	(204)
8.1 时间序列的对比分析	(204)
8.1.1 时间序列的含义	(204)

8.1.2	时间序列的速度分析	(205)
8.2	时间序列及其构成因素	(208)
8.2.1	时间序列的构成因素	(208)
8.2.2	时间序列构成因素的组合同模型	(210)
8.3	时间序列趋势变动分析	(211)
8.3.1	移动平均法	(211)
8.3.2	指数平滑法	(213)
8.3.3	线性趋势模型法	(215)
8.3.4	非线性趋势模型法	(217)
8.4	季节变动分析	(220)
8.4.1	季节变动分析的原始资料平均法	(220)
8.4.2	季节变动分析的趋势一循环剔除法	(222)
8.4.3	季节变动的调整	(223)
8.5	循环变动分析	(224)
8.5.1	循环变动及其测定目的	(224)
8.5.2	循环变动的测定方法	(224)
	本章小结	(228)
	思考与练习	(229)
	案例分析 中国社会消费品零售总额分析	(231)
第9章	统计指数	(233)
9.1	指数的概念与分类	(233)
9.1.1	指数的概念	(233)
9.1.2	指数的性质和分类	(233)
9.2	总指数的编制方法	(235)
9.2.1	总指数编制的基本问题	(235)
9.2.2	加权总指数的编制原理	(236)
9.2.3	加权综合指数的主要形式	(239)
9.2.4	加权平均指数的主要形式	(242)
9.3	指数体系与因素分析	(244)
9.3.1	指数体系及其作用	(244)
9.3.2	总量变动的因素分析	(245)
9.3.3	平均指标指数及其因素分析	(246)
9.4	几种常用的经济指数	(248)
9.4.1	生产指数和生产者价格指数	(248)
9.4.2	消费者价格指数和零售价格指数	(249)

9.4.3 股票价格指数	(249)
9.4.4 农副产品收购价格指数	(250)
9.4.5 产品成本指数	(251)
9.4.6 空间价格指数	(251)
9.5 综合评价指数	(252)
9.5.1 综合评价及其方法	(252)
9.5.2 构建综合评价指数的基本问题	(253)
9.5.3 综合评价指数的编制方法	(254)
本章小结	(257)
思考与练习	(259)
案例分析 统计指数及其分析	(261)
第 10 章 统计决策	(263)
10.1 统计决策的基本概念	(263)
10.1.1 统计决策的内涵	(263)
10.1.2 统计决策的基本步骤	(264)
10.1.3 收益矩阵表	(265)
10.2 完全不确定型决策	(266)
10.2.1 完全不确定型决策的准则	(266)
10.2.2 各种准则的特点和适用场合	(268)
10.3 一般风险型决策	(269)
10.3.1 自然状态概率分布的估计	(269)
10.3.2 风险型决策的准则	(270)
10.3.3 利用决策树进行风险型决策	(272)
10.4 贝叶斯决策	(274)
10.4.1 贝叶斯决策的内涵	(274)
10.4.2 贝叶斯公式与后验概率的估计	(274)
10.4.3 先验分析与后验分析	(275)
10.4.4 完全信息价值与补充信息价值	(276)
10.4.5 后验预分析	(279)
本章小结	(282)
思考与练习	(282)
案例分析 帮助王先生选择理财方案	(284)
第 11 章 国民经济统计	(286)
11.1 国民经济统计的内容和原则	(286)

11.1.1	国民经济运行与国民经济核算	(286)
11.1.2	国民经济核算体系	(286)
11.1.3	我国国民经济核算体系的基本框架	(287)
11.1.4	国民经济核算的原则	(288)
11.2	国民经济统计的分类	(290)
11.2.1	国民经济统计与国民经济分类	(290)
11.2.2	机构部门分类	(290)
11.2.3	活动部门分类	(291)
11.3	国民经济统计的基本指标	(292)
11.3.1	国民经济统计指标体系	(292)
11.3.2	国民经济生产指标	(292)
11.3.3	国民收入分配指标	(296)
11.3.4	对外经济往来指标	(298)
11.3.5	资产负债和国民财政指标	(300)
11.4	国民经济统计的常用指标	(302)
11.4.1	国民生产分析指标	(302)
11.4.2	收入分配分析指标	(303)
11.4.3	最终消费和投资分析指标	(304)
本章小结		(305)
思考与练习		(306)
案例分析 关于经济发展指标的选择		(308)
第 12 章	常用统计软件及其应用	(311)
12.1	统计软件的类型	(311)
12.1.1	统计软件的选择	(311)
12.1.2	常用统计软件的类型	(312)
12.2	SPSS 软件的简单应用	(314)
12.2.1	SPSS 软件的特点	(314)
12.2.2	利用 SPSS 进行统计处理的基本过程	(316)
12.2.3	SPSS 数据文件的建立	(317)
12.3	SAS 系统的简单应用	(321)
12.3.1	SAS 系统的特点	(321)
12.3.2	SAS 软件的操作	(322)
参考书目		(327)

第 1 章 统计学概论

1.1 统计数据与统计学

统计数据，即通过统计设计、统计调查、统计整理、统计分析而得到的有关社会经济现象的数据资料。在人们的日常生活和工作中，到处都有统计数据。例如，超市购物的客流量及货物进销统计；飞机场、列车站等旅客安检及登机、上车人数统计；学生考试成绩和名次统计；企业生产、销售、成本、利润和人员统计；国内生产总值及其增长率、物价指数、经理指数、制造业指数、股票指数等。

统计学的英文是“statistics”。它通常有两个含义：其单数名词表示“统计”；复数名词表示“统计学”、“统计数据”或“统计资料”。统计学是一门收集、整理、显示和分析统计数据的科学，其目的是探索数据内在的数量规律性。

那么，什么是数据内在的数量规律性呢？不妨用几个简单的例子来说明。

【例 1.1】 大家可能做过掷硬币和掷骰子的游戏，都知道随机地掷一次硬币或骰子不能事先确定出现正面还是反面或某个点数，也就是说个别游戏或试验中充满了不确定性或偶然性。机会游戏或赌博正是利用了这种不确定性和偶然性才能够吸引人。但当我们进行大量观察，即不断做重复试验时，就会发现掷一枚均匀硬币出现正面和反面的次数会大体相同，即比值接近 $1:1$ 。试验的次数越多，出现正面和反面的可能性就越接近 $1/2$ 这一稳定的数值。同样，在掷一枚均匀骰子时，出现 1、2、3、4、5、6 点数中任一点数的可能性也都接近 $1/6$ 。 $1/2$ 和 $1/6$ 就是掷硬币和掷骰子出现某一特定结果的概率，也就是我们探索出来的数量规律性。

【例 1.2】 农业科技人员在进行农作物试验时，如果其他试验条件相对固定，就会发现某种粮食作物的产量会随着某种施肥量的增加而增加。当开始增加施肥量时，产量增加较快。以后增加同样的施肥量，粮食产量的增加量逐渐减少。当施肥量增加到一定数值量，产量不再增加。这时如果再增加肥料，产量反而会减少。这一施肥量与粮食产量的数量关系就是我们探索的规律性。当他们从大量试验数据中用统计方法探索到施肥量与产量之间的数量关系，就可以考虑肥料的费用并选择最佳的施肥量了。

【例 1.3】 一般而言，随着广告费用的增加，企业的商品销售额会相应增加。对于不同的商品，投入相同的广告费用带来的销售额显然是不尽相同的。即使是同一种

商品,投入的第一个10万元广告费、第二个10万元广告费和第三个10万元广告费等所带来的销售增量也是不同的。运用统计方法就能从广告费和销售额的数据中找出其内在的数量规律,从而确定最佳广告策略。

为什么统计方法能够通过对数据的大量观察和处理而探索出其内在的数量规律性呢?这是由客观事物本身的特点和统计方法的特征共同决定的。从客观事物方面来说,根据辩证法的基本原理,任何客观事物都是必然性与偶然性的对立统一。同样,任何一个数据,也都是必然性与偶然性共同作用的结果,是二者作用的对立统一。必然性反映了事物本质的特征和联系,是比较稳定的,因而它决定了事物的内在本质是有规律可循的。应用统计方法就可以从偶然性探索到客观事物内在的本质的数量规律。

1.2 统计学的产生和发展

统计学产生于17世纪中叶,主要有以下三个来源。

统计学的第一个来源是英国威廉·配第(William Petty)的《政治算术》(1676年)。威廉·配第在书中用大量的数据对英国、法国、荷兰三国的经济实力进行比较,用数字、重量、尺度等定量的方法进行分析 and 比较,表达他的思想观点。马克思在《资本论》中评价威廉·配第“是政治经济学之父,在某种程度上也可以说是统计学的创始人”。

统计学的第二个来源是英国约翰·格朗特(John Graunt)的《关于死亡表的自然观察与政治观察》(1662年)。他通过大量观察的方法,研究并发现了人口与社会现象中重要的数量规律性。如新生儿的性别比率稳定在14:13;男性在各年龄组中死亡率高于女性;新生儿的死亡率较高;一般疾病与事故的死亡率较稳定而传染病的死亡率波动较大。在研究中,格朗特不但探索了人口变化和发展的一些数量规律,而且还对伦敦市总人口数量做出了较科学的估计。如果说配第是政府统计的创始人,那么,格朗特可被认为是人口统计的创始人。

统计学的第三个来源是法国的帕斯卡(B. Pascal)和费马(Pierre de Fermat)的古典概率论。早在他们之前,已有一些数学家在研究赌博中的数量规律了。意大利诗人但丁在15世纪就讨论过掷三颗骰子可能出现的各种点数。16世纪中,意大利科学家伽利略讨论了掷三颗骰子出现10点次数多于9点次数的原因。在数学家们对机会游戏研究的基础上,帕斯卡和费马通过通信的方式,将赌博中出现的各种具体问题,归纳为一般的概率原理,为后来概率论和统计学的发展奠定了重要的基础。

自17世纪中叶上述几位科学家从不同的角度开始了统计学研究后,经过几代统计学家的努力,历经两个半世纪,在19世纪末建成了古典统计学(主要是描述统计学)的基本框架。

20 世纪初,大工业的发展对产品质量检验问题提出了新的要求,即只抽取少量产品作为样本对全部产品的质量好坏做出推断。因为大批量产品要做全面的检验,既费时、费钱,又费人力,加之有些产品质量的检验要做破坏性检验,全部检验已不可能。1907 年,英国的戈塞特(W. S. Gosset)提出了小样本 t 统计量,利用 t 统计量就可以从大量的产品中只抽取较小的样本完成对全部产品质量的检验和推断,这样就使统计学进入了现代统计学(主要是推断统计学)的新阶段。以后经过著名统计学家费希尔(R. A. Fisher)给出了 F 统计量、最大似然估计、方差分析等方法和思想,奈曼(J. Neyman)和皮尔逊(E. S. Pearson)的置信区间估计和假设检验,沃尔德(A. Wald)的序贯抽样和统计决策函数等,到 20 世纪中叶构建了现代统计学的基本框架。

从 20 世纪 50 年代以来,统计学的理论、方法及其应用进入了一个全面发展的新阶段。一方面,统计学受计算机科学、信息论、混沌理论、人工智能等现代科学技术的影响,新的研究领域层出不穷,如多元统计分析、现代时间序列分析、贝叶斯统计、非参数统计、线性统计模型、探索性数据分析、数据挖掘等。另一方面,统计方法的应用领域不断扩展,几乎所有的科学研究都离不开统计方法。因为不论是自然科学、工程技术、农学、医学、军事科学,还是社会科学都离不开数据,要对数据进行研究和分析就必然要用到统计方法,现在连纯文科领域的法律、历史、语言、新闻等都越来越重视对统计数据进行分析,国外的人文与社会科学普遍开设统计学的课程,因而可以说统计学与数学、哲学一样成为所有学科的基础,并展现交叉性、边缘性、多科性、应用性、社会性、数量性、整体性等特点。

1.3 统计学的研究对象

统计学的研究对象,是指统计研究所要认识的客体,也就是客观事物总体的数量表现和数量关系。数量表现是指绝对数、相对数、平均数等。数量关系主要指相关关系、回归关系、指数关系等。

统计学的研究对象具有以下特征:

(1) 总体性:统计学的研究对象是以社会经济现象为主要客体,研究个体是为了更好地研究总体,从特殊到一般。总体离不开个体。

(2) 社会性:其研究对象不仅反映社会关系,而且反映生产关系和生产力水平。同时,统计工作是由许多统计工作者、各统计部门共同完成的,也具有一定的社会性和阶级性。

(3) 变异性:其研究对象随着客观事物的变化而变化,即使是同一指标或标志在不同时期的数值表现或特征也是不同的。统计上把总体各单位由于随机因素引起的标志表现的差异称为变异。

- (4) 具体性：其研究对象是实实在在的具体的社会经济现象。
- (5) 抽象性：其研究对象必然反映本质的规律性的东西，因而对个别事物研究之后要进行抽象的概括和总结，找出普遍性的规律。

1.4 统计学的分科

统计学的内容十分丰富，研究与应用的领域非常广泛。从统计教育的角度，统计学大致有两种分类：描述统计和推断统计；理论统计和应用统计。

1.4.1 描述统计和推断统计

描述统计 (descriptive statistics) 是用图形、表格和概括性的数字对数据进行描述的统计方法。图 1.1 的框图清楚地表明描述统计是对数据进行处理的第一阶段，即用直观的图形、汇总的表格和概括性的数字（如平均数）表示数据的分布、形状等特征，并为进一步的推断统计提供依据。

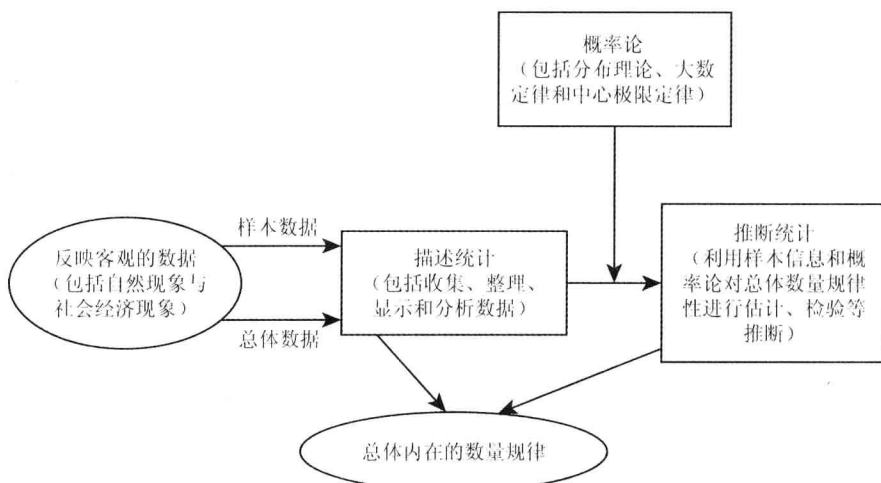


图 1.1 探索客观事物数量规律性的过程

推断统计 (inferential statistics) 是根据样本信息对总体进行估计、假设检验、预测或其他推断的统计方法。

统计学分为描述统计和推断统计，一方面反映了统计发展的前后两个阶段，另一方面反映了统计方法研究客观事物内在数量规律性的先后两个过程。

由图 1.1 可知，统计研究过程的起点是数据，终点是探索到客观事物总体内在的数量规律性。要达到统计研究的目的，如果我们收集到的是总体数据（如普查），则经

过描述统计之后就可以达到探索内在的数量规律性的目的了；但如果我们所获得的数据只是研究总体的一部分数据，要探索到总体的数量规律性，就必须应用概率论的理论，并根据样本整理出的信息对总体做出科学的推断。显然，描述统计是整个统计学的基础和统计研究工作的第一步，它包括对客观现象的度量，调查方案的设计，科学、及时、快速、经济地收集与整理数据，用图表显示数据，分析和提取数据中的有用信息以最终推断总体；推断统计是现代统计学的核心和统计研究工作的关键环节，因为统计最终能否科学准确地探索到总体内在的数量规律性与选用何种统计量，选用什么推断方法，如何进行推断有着直接的联系。

应该认识到，尽管描述统计可以在获得总体数据时直接探索出总体数量规律性，但这种情况在实际工作中很少见到。自然现象的总体多数是无限的，例如，统计物理研究中要弄清楚分子运动的规律，而分子优势无穷多的，不可能全部观察和试验。社会经济现象的总体虽然多数是有限的，但要考虑获得数据以及推断总体的时效性、经济性和准确性，抽样调查往往比普查更有效，因而应用也就更普遍。例如，全国的人口数量和变化、耕地面积、企业个数和经营情况等，虽然可以通过普查得到全部数据，但普查要投入大量的人力、财力和物力，而且学要很长的时间才能收集、整理出所要的数据，不是每年都能做的，因而我国确定每10年进行一次人口普查和经济普查，其他各年均以抽样调查数据进行判断。此外，大量的管理和研究工作不可能组织普查，例如，城市居民家庭每月的收入指出调查、某种商品的市场调查、某个事件的民意测验等都只能通过抽样调查方法，然后对总体数量规律性进行科学的推断。因而，在描述统计中收集、整理和分析的多是样本数据。这样，科学地整理样本数据、显示样本数据的特征和规律、提取样本数据中的有用信息就显得格外重要。

1.4.2 理论统计和应用统计

理论统计是指统计学的数学原理。从广义来讲，统计学是应该包括概率论的，因为概率论是统计推断的数学基础，而概率论是数学的一个分支，故理论统计是包括概率论在内的对统计方法数学原理的研究。

在统计工作者中，从事理论统计研究的人只是很少的一部分，大部分是应用统计方法去解决实际问题的应用统计工作者。统计学是一门数据科学，更强调应用统计。由于在自然科学、社会科学的所有研究和实际工作中都要通过数据来分析和解决问题，统计方法的应用自然而然地扩展到几乎所有的研究领域。

作为一名优秀的应用统计工作者，不但要能熟练地掌握和应用各种统计方法，而且必须具备所研究和应用领域的专业知识。因而，专业知识和统计方法是做好统计应用的两个基本功。这就要求统计应用人才是一名复合型人才，这也是人文、社科、理、工、农、医所有院校都开设统计学这门课程的缘由。