

统计系列教材

Introduction of Statistics

统计学导论

(第二版)

主编 向蓉美 王青华

西南财经大学出版社

统计系列教材

Introduction of Statistics

统计学导论

(第二版)

主编 向蓉美 王青华

西南财经大学出版社

图书在版编目(CIP)数据

统计学导论/向蓉美,王青华主编. —2版. —成都:西南财经大学出版社,
2008. 11
ISBN 978 - 7 - 81138 - 091 - 0

I. 统… II. ①向…②王… III. 统计学 IV. C8

中国版本图书馆 CIP 数据核字(2008)第 123855 号

统计学导论(第二版)

主编:向蓉美 王青华

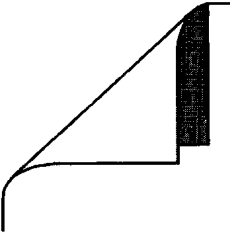
责任编辑:李雪

封面设计:杨红鹰

责任印制:封俊川

出版发行:	西南财经大学出版社(四川省成都市光华村街 55 号)
网 址:	http://www.xcpress.net
电子邮件:	xcpress@mail.sc.cninfo.net
邮政编码:	610074
电 话:	028 - 87353785 87352368
印 刷:	四川森林印务有限责任公司
成品尺寸:	170mm × 240mm
印 张:	14. 75
字 数:	270 千字
版 次:	2008 年 11 月第 2 版
印 次:	2008 年 11 月第 1 次印刷
印 数:	1—3000 册
书 号:	ISBN 978 - 7 - 81138 - 091 - 0
定 价:	25. 80 元

1. 如有印刷、装订等差错,可向本社营销部调换。
2. 版权所有,翻印必究。
3. 本书封底无本社数码防伪标志,不得销售。



第二版前言

《统计学导论》第一版出版以来，使用广泛，受到欢迎。为了使这本教材更加完善和便于组织教学，我们认真总结了教学实践的经验，由原作者对该教材进行了修订，内容有增有减。修订的主要内容有：

(1) 在市场经济条件下，统计调查中问卷设计越来越显示出它的重要性，本次修订加强了对问卷设计理论和方法的阐述。

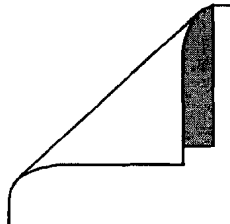
(2) 为了正确应用集中趋势和离中趋势指标，也为更好地理解后面章节的有关内容，本次修订增加了算术平均数和方差数学性质的内容。

(3) 编写这本教材的初衷主要是面向还没有学习过概率论的学生和读者使用的。根据一些教师和读者的意见，对第6章和第7章作了较大程度的修改。删除了第6章和第7章中有关概率论和假设检验原理的内容，使得教材内容更加简明、实用。抽样误差等概念和公式的符号也作了改动，以便与其他大部分教材所使用的符号保持一致，便于学生阅读其他参考文献。同时，在抽样估计部分，突出了实际应用中比较关心的抽样平均误差，增加了其他抽样组织方式的抽样估计等内容，更加有利于学生学以致用。

(4) 第二版中有的内容加了*，以方便不同层次、不同学时学生灵活选择。跳过这些内容不影响知识体系的相对完整和以后内容的学习。

(5) 全面更新了教材中例题和练习题的数据，有利于紧密联系实际进行教学。凡是我国国民经济实际数据，均取自2007年《中国统计年鉴》和中华人民共和国国家统计局2008年2月28日发布的《2007年国民经济和社会发展统计公报》。

编者
2008年7月



前言

当今的时代是信息的时代。统计学是一门研究数据的方法论学科，它能够为我们提供科学、可靠地收集信息的方式方法，教会我们如何“去粗取精、去伪存真”，从大量信息中找出现象的本质特征和变化规律，有助于我们利用手中的数据对复杂问题作出明智合理的决策。统计学应用范围很广，经济、社会、医学、生物、工程、技术等领域都有统计的广泛应用。正因为如此，统计学被教育部经济学和管理学教学指导委员会指定为经济类和管理类各专业必修的核心课程之一。

统计学的内容十分丰富。本书是关于统计学基本理论和方法的一本教材，适用于经济类和管理类非统计专业学生，也可以作为统计学专业的入门教材，也适用于其他专业学生和广大经济管理人员作为学习统计学的参考书。

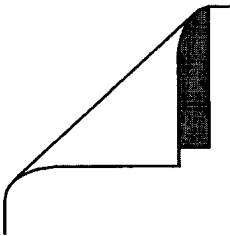
在多年的教学实践中，我们发现很多学生因对数学语言和公式推导的畏惧而对统计学失去信心，因对大量的计算感到乏味而对统计学失去兴趣。我们认为，随着计算机和统计软件的普及，大量的计算工作完全可以交给计算机去完成，学生学习统计是为了掌握统计思想，把统计作为一门实用的工具。因此，在本教材的编写过程中，我们强调“重思想而轻公式”、“重应用而轻推导”、“重思考而轻演算”。在阐述统计理论方法时，深入浅出，尽可能利用通俗易懂的语言进行解释，避免不必要的公式推导和数据演算，把统计知识的学习与 *Excel* 软件运用有机结合，注重教会学生思考如何收集所需数据、应该采用什么方法去整理和分析数据以及怎样利用计算机软件输出的结果得出结论、进行解释，注重培养学生应用统计知识去分析和解决实际问题的能力。每章后面配有本章小结以及思考与练习题，便于把握本章的学习要点，并通过思考与练习题加深对本章内容的理解和掌握。

本教材由西南财经大学统计学院几位长期从事统计学教学的教授和副教授共同编写而成。其中，向蓉美，第1章和第10章；徐浪，第2章和第9章；雷敏，第3章；苏远琳，第4章；聂富强，第5章；王青华，第6、7、8章和第9章第4节。最后由向蓉美和王青华对全书进行了总纂。

随着社会经济的不断发展和统计理论的不断完善，我们将适时根据其发展变化情况修订本教材，也恳请读者将使用本教材的建议和意见及时反馈给我们，对此我们表示衷心的感谢。

作者

2006 年 6 月于西南财经大学



目 录

1	总论	1
§ 1.1	什么是统计	1
§ 1.1.1	统计的涵义	1
§ 1.1.2	统计研究的基本方法	5
§ 1.1.3	统计研究的全过程	6
§ 1.2	统计中的基本概念	7
§ 1.2.1	总体、个体与样本	7
§ 1.2.2	标志、指标与变量	9
§ 1.2.3	统计数据	11
§ 1.2.4	参数与统计量	13
§ 1.3	用 <i>Excel</i> 进行统计分析	13
§ 1.3.1	<i>Excel</i> 的工作界面	14
§ 1.3.2	<i>Excel</i> 的有关统计功能	15
	本章小结	17
	思考题与练习题	18
2	统计数据的收集	19
§ 2.1	统计调查的基本要求和调查方案	19
§ 2.1.1	统计调查的意义及要求	19
§ 2.1.2	统计误差	20
§ 2.1.3	统计调查方案	21
§ 2.2	原始数据的收集	23
§ 2.2.1	统计调查方式	23
§ 2.2.2	原始资料收集的具体方法	27
§ 2.2.3	问卷设计	28

§ 2.3 次级资料的收集 31

§ 2.3.1 次级资料的来源 31

§ 2.3.2 次级资料的局限性 32

本章小结 33

思考题与练习题 34

3 统计数据的整理 35

§ 3.1 统计分组 35

§ 3.1.1 统计分组的概念和作用 35

§ 3.1.2 统计分组的原则 36

§ 3.1.3 统计分组的类型 36

§ 3.2 分布数列 37

§ 3.2.1 分布数列的概念和种类 37

§ 3.2.2 变量数列的编制 40

§ 3.2.3 次数分布图 42

§ 3.2.4 累计次数分布与洛伦茨曲线 46

§ 3.3 统计图与统计表 48

§ 3.3.1 统计图 48

§ 3.3.2 统计表 53

本章小结 55

思考题与练习题 55

4 总量指标及对比分析 58

§ 4.1 总量指标 58

§ 4.1.1 总量指标的意义 58

§ 4.1.2 总量指标的计量单位 59

§ 4.1.3 总量指标的时间属性 59

§ 4.2 对比分析 60

§ 4.2.1 对比分析的意义 60

§ 4.2.2 常用的对比分析方法 62

§ 4.2.3 对比分析的原则 67

本章小结 68

思考题与练习题 69

- 5 数据分布特征的描述 71
 - § 5.1 集中趋势的测度 71
 - § 5.1.1 数值平均数 71
 - § 5.1.2 位置平均数 77
 - § 5.1.3 位置平均数与算术平均数的比较 80
 - § 5.1.4 计算和应用平均指标应注意的问题 81
 - § 5.2 离中趋势的测度 81
 - § 5.2.1 极差与四分位差 82
 - § 5.2.2 平均差、标准差与方差 82
 - § 5.2.3 离散系数 84
 - § 5.3 数据分布形态的测定 85
 - § 5.3.1 偏态的测定 85
 - § 5.3.2 峰态的测定 87
 - § 5.4 利用 Excel 计算数据分布特征指标 88
- 本章小结 90
- 思考题与练习题 91
- 6 抽样估计 95
 - § 6.1 抽样估计的基本问题 95
 - § 6.1.1 总体参数与估计量 95
 - § 6.1.2 估计量的评选标准 96
 - § 6.1.3 抽样误差的测定 98
 - § 6.1.4 参数估计的方法 101
 - § 6.2 总体均值和成数的区间估计 101
 - § 6.2.1 大样本条件下总体均值的区间估计 101
 - * § 6.2.2 小样本条件下正态总体均值的区间估计 104
 - § 6.2.3 大样本条件下总体成数的区间估计 105
 - § 6.3 抽样数目的确定 106
 - § 6.3.1 估计总体均值所必要的抽样数目 106
 - § 6.3.2 估计总体成数所必要的抽样数目 108
 - * § 6.4 其他抽样组织方式的抽样估计 108
 - § 6.4.1 分层抽样及其抽样估计 109
 - § 6.4.2 等距抽样及其抽样估计 111

§ 6.4.3 整群抽样 112

本章小结 114

思考题与练习题 115

7 相关与回归分析 118

§ 7.1 相关关系分析的概述 118

§ 7.1.1 相关关系的概念 118

§ 7.1.2 相关关系的类型 119

§ 7.1.3 相关关系分析的基本内容 120

§ 7.2 一元线性相关分析 121

§ 7.2.1 散点图 121

§ 7.2.2 相关系数 122

§ 7.3 一元线性回归分析 125

§ 7.3.1 一元线性回归方程的估计 125

§ 7.3.2 一元线性回归方程的拟合效果 128

§ 7.3.3 利用一元线性回归方程进行预测 130

本章小结 133

思考题与练习题 133

8 时间序列分析 138

§ 8.1 时间序列概述 138

§ 8.1.1 时间序列的概念和种类 138

§ 8.1.2 时间序列的构成与分解 140

§ 8.1.3 时间序列的编制原则 142

§ 8.2 时间序列的水平与速度分析 143

§ 8.2.1 时间序列的水平分析 143

§ 8.2.2 时间序列分析的速度指标 147

§ 8.2.3 水平分析与速度分析的结合与应用 152

§ 8.3 时间序列的趋势测定与预测 153

§ 8.3.2 移动平均法 153

§ 8.3.3 指数平滑法 156

§ 8.3.4 趋势方程拟合法 160

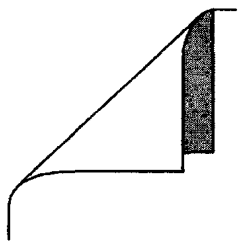
§ 8.4 时间序列的季节变动测定 164

§ 8.4.1	测定季节变动的同期平均法	165
§ 8.4.2	测定季节变动的移动平均趋势剔除法	167
§ 8.4.3	季节影响的调整	170
§ 8.5	时间序列的循环变动和不规则变动的测定	171
§ 8.5.1	循环变动的测定方法	171
§ 8.5.2	不规则变动的测定	173
本章小结		174
思考题与练习题		175
9 统计指数 178		
§ 9.1	统计指数的概念、作用和分类	178
§ 9.1.1	统计指数的概念	178
§ 9.1.2	统计指数的作用	180
§ 9.1.3	统计指数的分类	180
§ 9.2	总指数的计算	181
§ 9.2.1	综合法指数	182
§ 9.2.2	平均法指数	186
§ 9.3	指数体系与因素分析	189
§ 9.3.1	指数体系的概念和作用	189
§ 9.3.2	现象总量变动的因素分析	190
§ 9.3.3	总平均数变动的因素分析	193
§ 9.3.4	含总平均数的现象总量变动因素分析	195
§ 9.4	实际经济生活中的几种重要指数	196
§ 9.4.1	工业生产指数	196
§ 9.4.2	居民消费价格指数	197
§ 9.4.3	股票价格指数	200
本章小结		201
思考题与练习题		202
10 综合评价 204		
§ 10.1	综合评价概述	204
§ 10.1.1	综合评价的意义	204
§ 10.1.2	综合评价的一般步骤	205

§ 10.1.3	综合评价方法的局限性	206
§ 10.2	评价指标的选择与数据处理	207
§ 10.2.1	评价指标的选择原则	207
§ 10.2.2	评价指标的同向化和同度量化方法	208
§ 10.2.3	指标赋权方法	209
§ 10.3	综合评价方法	213
§ 10.3.1	计分法	213
§ 10.3.2	加权指数法	216
§ 10.3.3	改进功效系数法	219
本章小结		220
思考题与练习题		220

附录 223

附表 1	标准正态分布函数值表	223
附表 2	t 分布上侧分位数表	225



1 总论

有人说统计不仅是统计人谋生的手段,也是现代人谋生的手段,因为我们生活在信息时代,信息的主体就是统计数据。统计的力量如此强大、无处不在,凡是有数据的地方,就有统计。好、比较好、很好、非常好、比以往任何时候都好等这样的词语太苍白、太空洞了;而 2007 年我国国内生产总值达到 246 619 亿元,按可比价比 1978 年增长 12 倍多,这样的统计数据实实在在地反映了中国改革开放以来取得的巨大成就。

什么是统计?统计能做什么?如何进行统计?有些什么样的统计数据?在这一章中,我们将讨论这些统计学的基本问题。

§ 1.1 什么是统计

§ 1.1.1 统计的涵义

1.1.1.1 感悟统计

无论我们是否学过统计、懂得统计,我们生活中的每一天都会遇到大量统计问题,新闻和大众媒体每天都在表现统计数字。例如,每年净增加人口数据;每年、季、月度的增加值及其增长速度数据;每年的《政府工作报告》总是要列举大量的水平、比例、结构、速度等数据,说明国民经济的发展状况;统计部门每月、每年都要公布价格变动数据,说明与人民生活和社会生产息息相关的价格状况等。

宇宙间万事万物,林林总总,各种事物可依特定的性质予以归类,形成各种群体——现象总体。无论是属于自然的、实验的,还是社会的、经济的,凡是可以数据表现的总体,都可以作为统计的研究对象。统计方法和统计思想渗透到了社会、经济、自然、科技、生活的每一个角落,统计学不仅在社会经济领域得到发展,而且一

些过去与数量毫无联系的学科,如政治学、法学、历史学、艺术学、考古学等都在对应用统计方法进行研究和实践。

1.1.1.2 统计的涵义

“统计”作为社会经济生活中经常使用的名词,一般具有三种含义:统计工作、统计资料和统计科学。

(1) 统计工作

统计工作即统计实践活动,是人们为了说明所研究对象的某种数量特征和数量规律性,而对该现象的数据进行收集、整理与分析的活动。例如,为了获得粮食产量而进行的抽样调查活动,为了获得全国人口的数量和构成等而进行的人口普查活动等。

统计实践活动有着悠久的历史,早在公元前 2000 多年,统治者为了征兵、征税、管理奴隶的需要,就有了人口、土地、财产的统计活动。《通典》记载了我们历史上最早的统计数据,“禹平水土,为九州”、“九州之地凡 24 388 024 顷,人口 13 553 923 人”。到中世纪,西欧各国都有了人口、军队、领地、财产等的统计活动。

(2) 统计资料

统计资料即统计数据,是通过统计工作所获得的能够说明现象总体某种特征的数据,是统计实践活动的成果。例如 2007 年我国国内生产总值 246 619 亿元,按可比价计算比 2006 年增长 11.4% 等,就是广大统计部门和统计工作者辛勤劳动得到的我国经济发展水平的统计资料。

统计数据最集中、最系统地反映在各种统计年鉴中,如《中国统计年鉴》、《国际统计年鉴》、《中国经济年鉴》、《中国金融年鉴》、《中国物价及城镇居民家庭收支调查统计年鉴》等。《中国统计年鉴》涵盖了国民经济所有主要数据,包括国民经济核算、各行业生产、就业、人民生活、对内对外贸易、社会活动、环境等方面的数据。随着互联网技术的发展和普及,在网站上查找数据有成为统计数据主要来源的趋势。统计数据还可以从经济分析专著、期刊里获得。比如《中国经济数据分析》、《经济形势分析与预测(经济蓝皮书)》,一些期刊刊登的最新月度数据等。

(3) 统计学

统计学是一门研究收集数据、表现数据、分析数据、解释数据,从而认识现象数量规律的方法论科学。统计学源于统计实践活动,是对统计实践活动的理论概括和总结,又用于指导统计实践活动。

收集数据需要对客观现象作周密细致的调查;表现数据需要对调查得到的数据加以整理,使之成为反映现象总体的条理化、系统化的数据;分析数据需要用科学的方法从数据中得出反映现象本质的数量规律性;解释数据需要用有关知识对数量规

律性的原因作出说明,所有这些就构成了统计学的研究内容。

统计学是一门为定量分析提供方法的方法论科学,因而是一门应用性很强的科学。它与几乎所有的学科领域都有着或多或少的联系,凡是有数据的地方,就有统计学的用武之地。统计学为其他学科提供研究数量规律性的方法,但是统计学绝不是万能的,各个学科数量规律的解释还需要由各学科的理论来完成。比如,大量观察发现我国人口的性别比随着年龄的增大呈下降趋势。形成这样的比例和趋势的原因,不是统计学能够解释的,要用遗传学、医学甚至社会学的理论来解释。又如大量观察表明,吸烟者患肺癌的比例大于不吸烟者患肺癌的比例,吸烟是否会导致患肺癌?为什么会导致患肺癌?这是医学研究的问题。

尽管统计学不能解决各学科的所有问题,但是,各门学科离不开统计学,统计学的理论和方法在各门学科的研究中会发挥越来越重要的作用。

作为一门科学,统计学初创于17世纪中叶至18世纪初,当时主要的学派有国势学派和政治算术学派。

国势学派的创始人是德国人赫尔曼·康令。康令1660年以后,开始定期地、系统地用对比的方法讲授国家比较方面的知识,他不仅讲述事实,而且试图探讨事实的因果关系。他把这个课程叫做《欧洲最近国势学》,“国势学”由此产生。当时,康令的学说产生了很大的影响,德国大学中的许多教授都称赞并追随他的学说。其中戈特弗里德·阿亨瓦尔发扬了康令的学术思想,把这门课程定名为“统计学”,开始有了“统计学”这个名称。但是国势学派对各国情况作一般性的比较记述,如“某国人口众多”、“土地辽阔”之类,而没有进行数量研究和描述。所以国势学派是有统计学之名,而无统计学之实。

政治算术学派的创始人是英国人威廉·配第。配第首先提出了用数量科学地研究社会经济现象的方法——政治算术。在他的名著《政治算术》序言中,配第明确指出:“我进行这项工作所使用的方法,在目前还不是常见的。因为和只使用比较级或最高级的词语以及单纯作思维的论证相反,我却采用了这样的方法(作为我很久以来就想建立的政治算术的一个范例),即用数字、重量和尺度的词汇来表达我自己想说的内容,只进行诉诸人们感官的论证和考察在性质上有可见的根据的原因。”他还说:“用数字、重量和尺度(它们构成我下面立论的基础)来表示的展望和论旨,都是真实的,即使不真实,也不会有明显的错误。……‘因为,能够证明为确实的东西,也就是确实的。’”他在研究社会经济现象的规律时,还应用推算法、分组法,编制原始的图表,计算一系列的总量指标、相对指标和平均指标。但是配第始终没有用“统计学”三个字,所以政治算术学派是有统计学之实,而无统计学之名。

近代统计学产生于18世纪末至19世纪末,主要的学派是数理统计学派。数理

统计学派的代表人物是比利时著名的统计学家、数学家、物理学家、天文学家和人类学家兰贝特·阿道夫·雅克·凯特勒。他融会贯通各家各派的统计思想,博采众长,把统计学发展中的三个主要源泉,即德国的国势学、英国的政治算术和意大利、法国的古典概率论,加以协调、统一、改造和融合成为具有近代意义的统计学,把统计学推向了新的发展阶段。他将统计方法用于研究人类,促进了人口统计学的发展;他提出“平均人”理论,用平均值作为实际值的一种代表值;他创建“国际统计会议”组织,促进了国际统计交流与合作。可以说凯特勒是古典统计学的终结者、近代统计学的先驱者,在统计学发展史上起着承先启后、继往开来的作用。

现代统计学一般认为从 19 世纪末 20 世纪初开始,其标志是推断统计学的问世。英国生物学家、统计学家卡尔·皮尔森在 1894—1895 年提出包括正态分布、矩形分布、J 型分布、U 型分布等 13 种曲线及其方程式。他的这一研究成果,打破了以往次数分布曲线的“唯正态”观念,推进了次数分布曲线理论的发展和应用,为大样本理论奠定了基础;他还提出了卡方检验和标准差概念,发展了相关回归理论等。英国统计学家威廉·西利·戈塞特创立了 t 分布,开创了小样本理论的先河。

统计学的发展史表明,随着社会的发展与实践的需要,统计学越来越多地依赖和吸收数学方法,使统计方法不断丰富和完善,使统计学不断发展和演变,成为研究社会经济现象、自然技术现象数量方面的方法论科学。目前,统计学越来越多地向其他学科领域渗透,形成各种以统计学为基础的边缘学科,随着统计学应用日益广泛和深入,特别是借助电子计算机,统计学所发挥的功效必将日益增强。

1.1.1.3 统计的类型

统计可以分为描述统计和推断统计,相应的统计学也可以分为描述统计学和推断统计学。

用从一个总体或样本中收集到的数据,来对这个总体或样本进行描述或得出有关这个总体或样本的结论,这样的统计称为描述统计。如全班有 50 名同学或从全校 15 000 名同学中抽取 400 名同学,把这 50 名或 400 名同学的成绩用图、表或特征值如平均分数、分数的标准差、及格率等表示出来,从而得出针对该班或 400 名同学学习情况的结论。描述统计学是关于如何对现象的数据特征进行观测、整理、计量、表述的理论和方法论科学。

用从一个总体中随机抽取的样本数据,得出关于这个总体的结论,这样的统计称为推断统计。如从全校 15 000 名学生中,随机抽取 400 名学生进行学习情况调查。用这 400 名学生平均分数、分数的标准差、及格率等推算出全校 15 000 名学生的平均分数、分数的标准差、及格率等,从而得出全校学生学习情况的结论,这就是推断统计。推断统计学是关于如何抽取样本并利用样本数据推断总体有关数据的

理论和方法论科学。推断统计主要包括参数估计和假设检验两种类型。

只有对样本进行了描述,才能对总体进行推断,所以描述统计是推断统计是基础。本教材第二章至第五章属于描述统计学的基本理论和方法,第六章属于推断统计学的基本理论和方法。

1.1.1.4 统计的特点

统计具有下列两个鲜明的特点:

(1) 数量性

统计的数量性特点指统计总是用数字作为语言来表述事实。统计运用科学的方法收集、表现、分析和解释数据,并用特有的统计指标表明所研究现象的规模、水平、比率、依存度、发展变化趋势和规律等。

但是统计不研究抽象的数量,它是在质的规定性下研究数量。例如,只有明确了国内生产总值的概念和范围,才能得到国内生产总值数据;只有规定了工资总额的内涵和外延,才能得到工资总额数据。

(2) 总体性

统计的总体性特点指统计不是研究个别事物的数据,而是研究大量个别事物构成的现象整体的数据,只有这样才能达到认识现象的数量规律的目的。例如,劳动力资源统计,不是要了解个别劳动力的情况,而是要反映一个国家或地区的劳动力资源总数及其构成、就业总数及其分布等。

但是,统计是从认识个别事物入手来认识现象整体的数量特征的。例如要了解劳动力资源总数及其构成、就业总数及其分布等,必须从每一个别劳动力开始,对其性别、年龄、文化程度、职业状况进行调查登记,然后经过分类汇总计算,才能了解劳动力资源的整体情况。

§ 1.1.2 统计研究的基本方法

统计研究的方法多种多样,最基本的方法有:

1.1.2.1 大量观察法

所谓大量观察法是指为了对现象整体的数量规律有所了解,必须对所研究现象的全部或足够多的个体进行调查的方法。只有通过大量观察法才能消除偶然的、次要的因素的影响,以反映主要的、共同起作用的因素所呈现的规律性,达到对现象总体数量规律的认识。例如,就个别家庭来说,可能有的男性人口多些,有的女性人口多些,但是如果观察上千上万个家庭,就会发现人口的性别比例大约为 1:1;又如掷硬币或掷骰子,每掷一次出现哪一面或哪个点子是不确定的,但是当我们掷很多次硬币或骰子时,就发现掷硬币出现正面和反面的可能性几乎各为 50%,掷骰子出