

21世纪经济管理精品教材·经济学系列

统计学：原理及应用

孔 锐 高孝伟 何大义 韩丽红 编著



清华大学出版社

21世纪经济管理精品教材·经济学系列

统计学：原理及应用

孔 锐 高孝伟 何大义 韩丽红 编著

清华大学出版社
北 京

内 容 简 介

本书具有以下三个主要特点：一是根据统计学教学大纲，兼顾了教学内容的系统性和完整性，不仅对统计学基本原理和基本方法进行了比较全面的阐述，而且构建了注重理论与实际相结合的知识应用系统，包括知识点应用案例学习层、知识点练习学习层和知识综合应用案例学习层。二是为各类人员的学习提供了方便。包括每章后归纳了关键知识点，备有训练用题和参考答案。三是结合本教材的内容，介绍了最常用系统软件使用方法和步骤，为学生提高处理大量统计数据的能力奠定了基础。

本书以国内外统计学文献为参考，结合社会经济发展的实际，紧密围绕经济管理专业教学改革的需要，围绕相关课程教学大纲，尽可能地采用实际的社会经济发展数据，增加了统计案例分析的内容或比重，使得即使在一些原理并不能完全理解的情况下，学生也能够运用统计分析软件和工具进行必要的统计分析。

本书适合于非统计学专业的研究生、MBA 和本科生的统计课程使用，也可以作为有关专业人员培训和自学的参考教材。

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。

版权所有，侵权必究。侵权举报电话：010-62782989 13701121933

图书在版编目(CIP)数据

统计学：原理及应用/孔锐等编著. --北京：清华大学出版社，2016

(21 世纪经济管理精品教材·经济学系列)

ISBN 978-7-302-44189-2

I. ①统… II. ①孔… 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(2016)第 174910 号

责任编辑：刘志彬

封面设计：汉风唐韵

责任校对：宋玉莲

责任印制：何 芊

出版发行：清华大学出版社

网 址：<http://www.tup.com.cn>, <http://www.wqbook.com>

地 址：北京清华大学学研大厦 A 座 邮 编：100084

社 总 机：010-62770175 邮 购：010-62786544

投稿与读者服务：010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈：010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者：三河市吉祥印务有限公司

经 销：全国新华书店

开 本：185mm×260mm

印 张：23.75

字 数：544 千字

版 次：2016 年 8 月第 1 版

印 次：2016 年 8 月第 1 次印刷

印 数：1~4000

定 价：45.00 元

产品编号：068905-01



前言

统计学是一门认识论和方法论学科,是使用数学和其他学科的专业知识进行数据收集、整理、分析研究对象的本质的一门综合性科学。统计分析是经济管理活动中经常用到的定量研究工具,其独特的研究对象和方法决定了它能够成为指导科学研究和宏观社会经济管理的指示器。统计学不仅是从事自然科学和社会科学研究的人所必须掌握的一门科学,也是从事社会活动的人,尤其是从事管理工作的人应该学习并掌握的一门科学。因此,统计学不仅是经济、管理类专业的学生在校期间系统学习的主要课程之一,其分析方法和工具也应该为其他专业学生所了解和掌握。

在进行“高等学校‘专业综合改革试点’项目”的研究过程中发现,已经出版的各种统计学教材虽然很多,但随着大学人才培养定位从精英式教育到通才式培养、拓宽基础教育的转变,统计学教学也必须顺应而变革,寻找更加明确并且差异化的教学需求定位。尤其是现代办公工具和统计软件的广泛使用和普及,使得相关教育者感到传统的统计学教材内容需要进行调整、补充和完善,以进一步适应社会发展的需要。根据非统计专业学生学习的特点和要求,本书在系统介绍统计学基本原理的基础上,大量使用社会经济发展的真实数据和案例展开分析,以便读者加深对统计学原理的理解并充分掌握统计分析方法的使用。

本书以国内外统计学文献为参考,结合社会经济发展的实际,紧密围绕经济管理专业教学改革的需要,围绕相关课程教学大纲,尽可能地采用实际的社会经济发展数据,增加了统计案例分析的内容或比重,使得即使在一些原理并不能完全理解的情况下,学生也能够运用统计分析软件和工具进行必要的统计分析。本书不仅适合于非统计学专业的研究生、MBA 和本科生的统计课程使用,也可以作为有关专业人员培训和自学的参考教材。

本书具有以下三个主要特点:一是根据统计学教学大纲,兼顾了教学内容的系统性和完整性,不仅对统计学基本原理和基本方法进行了比较全面的阐述,而且构建了注重理论与实际相结合的知识应用系统,包括知识点应用案例学习层、知识点练习学习层和知识综合应用案例学习层。二是为各类人员的学习提供了方便。包括每章后归纳了关键知识点,备有训练用题和参考答案。三是结合本书的内容,介绍了最常用系统软件使用方法和步骤,为学

生提高处理大量统计数据的能力奠定了基础。

本书的编写得到了教育部“高等学校‘专业综合改革试点’项目”的支持。由中国地质大学(北京)的多位具有扎实理论基础和丰富教学经验的教师合作编写。全书由孔锐、高孝伟主要执笔。第1、6、9章和10章由高孝伟执笔,第2章由韩丽红执笔,第3、4章由孔锐执笔,第5、7章由何大义执笔,第8章由徐春骐和孔锐执笔,第11章由何大义和孔锐执笔。全书由孔锐和高孝伟统稿和审稿。在资料查询、收集工作中薛芳芳、郭晓飞、翟海燕、张思梦、赖丽娟、韩欣怡、杨贺林和王文通等研究生做了大量的具体工作。

由于作者水平有限,不足之处在所难免,请读者给予批评指正。

编 者

2015年11月



目录

第一章 统计学概论	1
第一节 统计与统计学	1
一、统计学的产生和发展	1
二、统计学的性质和特点	2
三、统计学的分科	4
第二节 统计学的方法及理论基础	5
一、统计学的基本方法	5
二、统计学的理论基础	6
第三节 统计学的基本概念	8
一、统计总体和总体单位	8
二、标志和标志表现	9
三、统计指标和统计指标体系	10
第四节 统计指标的表现形式	12
一、统计指标的分类	12
二、总量指标	13
三、相对指标	14
四、平均指标	16
【附】	16
实证案例分析	16
本章关键知识点	19
复习思考题	19
练习题	20
轻松一刻	21
参考文献	21
第二章 数据资料的搜集与整理	22
第一节 数据的意义及类型	22
一、数据的意义	22

二、数据的计量与类型	22
三、统计数据的质量	25
第二节 数据的搜集与整理	26
一、数据搜集方式的基本类型	27
二、原始数据搜集的主要方式	28
三、数据搜集的方法	32
四、统计资料的整理	35
第三节 统计调查方案的设计	38
一、统计调查方案的内容	38
二、统计调查的组织和实施	40
第四节 调查表和调查问卷	40
一、调查表	40
二、调查问卷	41
【附】	46
实证案例分析	46
本章关键知识点	50
复习思考题	50
练习题	51
轻松一刻	52
参考文献	53
第三章 统计分组与频数分布	54
第一节 统计分组	54
一、统计分组的概念和作用	54
二、统计分组标志选择原则	56
三、统计分组的方法	57
第二节 频数分布	60
一、频数分布的概念和种类	60
二、频数分布数列的编制程序	62
三、频数分布的表示方法	65
四、洛伦茨曲线	69
第三节 统计表	70
一、统计表的构成和概念	70
二、统计表的类型	71
三、统计表的设计原则	73
【附】	74
实证案例分析	74
本章关键知识点	77

复习思考题	77
练习题	77
轻松一刻	80
参考文献	81
第四章 集中趋势和离散程度的确定	82
第一节 集中趋势和离散程度	82
一、总体的集中趋势	82
二、总体的离散程度	82
三、偏态与峰态	83
第二节 算术平均数	84
一、算术平均数的计算	84
二、加权算术平均数的特性	86
三、算术平均数的数学性质及应用	87
四、算术平均数应用注意事项	89
第三节 调和平均数	90
一、调和平均数的计算	90
二、调和平均数的特点	90
三、调和平均数的应用	91
第四节 几何平均数	92
一、几何平均数的计算	92
二、几何平均数的特点和应用	92
第五节 中位数和众数	93
一、中位数	93
二、分位数	96
三、众数	96
四、算术平均数、中位数、众数之间的关系	97
第六节 标志变异指标	98
一、极差	98
二、平均差	99
三、标准差	100
四、交替标志	101
五、方差	103
六、离散系数	104
【附】	105
实证案例分析	105
关键知识点	107
复习思考题	107

练习题	108
轻松一刻	111
参考文献	111
第五章 抽样推断与假设检验	112
第一节 抽样推断	112
一、抽样推断的基本概念	112
二、抽样分布	117
三、抽样推断的基本方法	123
四、样本容量的确定	131
第二节 假设检验	133
一、假设检验的基本问题	133
二、总体均值的假设检验	137
三、总体比例的假设检验	141
四、总体方差的假设检验	142
【附】	144
实证案例分析	144
关键知识点	145
复习思考题	145
练习题	145
轻松一刻	147
参考文献	147
第六章 列联分析	148
第一节 分类数据与列联表	148
一、分类数据	148
二、列联表的构造	148
三、列联的分布	149
第二节 拟合优度检验	150
一、 χ^2 统计量	150
二、拟合优度检验	151
第三节 独立性检验	152
一、问题的提出	152
二、独立性检验举例	153
第四节 列联表中的相关测量	154
一、 ϕ 相关系数	154
二、列联相关系数 c	155
三、 V 相关系数	155

四、相关系数数值分析	155
【附】	156
实证案例分析	156
本章关键知识点	159
复习思考题	159
练习题	159
轻松一刻	160
参考文献	161
第七章 方差分析	162
第一节 方差分析的基本思想	162
一、方差分析的相关概念	162
二、方差分析的基本原理	163
第二节 单因素方差分析	164
一、进行单因素方差分析需要计算的数据	164
二、单因素方差分析的步骤	166
三、关系强度的测量	167
第三节 双因素方差分析	168
一、双因素方差分析的数据结构	168
二、离差平方和的构成	169
【附】	171
实证案例分析	171
本章关键知识点	172
复习思考题	172
练习题	172
轻松一刻	173
参考文献	174
第八章 相关与回归分析	175
第一节 变量间的相关关系	175
一、相关的概念和种类	175
二、相关分析与相关系数	177
第二节 一元线性回归分析	181
一、回归分析的含义	181
二、相关分析与回归分析的区别和联系	182
三、一元线性回归方程模型	183
四、回归直线的拟合优度	186
五、显著性检验	188
第三节 多元线性回归分析	190

一、多元线性回归模型	190
二、统计显著性检验	191
第四节 可化为线性回归分析的曲线回归	192
一、指数函数	193
二、幂函数	193
三、双曲线函数	193
四、对数函数	193
五、S 形曲线	193
【附】	193
实证案例分析	193
本章关键知识点	197
复习思考题	198
练习题	199
轻松一刻	201
参考文献	201
第九章 时间序列分析	202
第一节 时间序列及其分析指标	202
一、时间序列的概念和种类	202
二、时间序列的分析指标	203
第二节 长期趋势分析	210
一、长期趋势的概念	210
二、长期趋势分析方法	210
第三节 季节变动和循环变动分析	221
一、季节变动分析	221
二、循环变动分析	222
【附】	225
实证案例分析	225
本章关键知识点	227
复习思考题	227
练习题	227
轻松一刻钟	230
参考文献	230
第十章 统计指数	231
第一节 统计指数的概念、分类与作用	231
一、统计指数的概念	231
二、统计指数的分类	231
三、统计指数的性质和作用	233
第二节 统计指数的编制	233

一、个体指数的编制	233
二、总指数的编制	234
三、其他形式的综合指数	239
第三节 总量指标指数体系及因素分析	242
一、总量指标指数及指数体系	242
二、总量指标指数因素分析的方法	243
第四节 平均指标指数体系及因素分析	248
一、平均指标指数及指数体系	248
二、平均指标指数因素分析	250
第五节 两种常见的价格指数	251
一、商品零售价格指数	251
二、股票价格指数	253
【附】	254
实证案例分析	254
本章关键知识点	256
复习思考题	256
练习题	256
轻松一刻	258
参考文献	259
第十一章 统计软件的应用	260
第一节 Excel 在统计学中的运用	260
一、数据输入	260
二、描述统计	262
三、区间估计	270
四、假设检验	271
五、单因素方差	275
六、相关系数计算	276
七、线性回归分析	277
第二节 SPSS 软件在统计学中的运用	278
一、数据输入	279
二、描述统计	282
三、相关系数计算	301
四、线性回归分析	304
五、时间数列分析	311
【附】	321
轻松一刻	321
参考文献	321

附录 1 常用统计表	322
表 1 标准正态分布表	322
表 2 t 分布表	324
表 3 χ^2 分布表	325
表 4 F 分布表	327
附录 2 练习题答案	331
第一章	331
第二章	332
第三章	335
第四章	341
第五章	344
第六章	350
第七章	351
第八章	353
第九章	355
第十章	359

统计学概论

第一章是以后各章节的基础,主要介绍统计学的产生与发展、统计学性质和特研究对象、统计研究的主要方法及理论基础,统计研究所涉及的基础性概念、统计研究和实践中常用的指标及指标体系的概念、统计指标的形成和三种表现形式。

第一节 统计与统计学

一、统计学的产生和发展

统计作为一种社会实践活动,大约产生于奴隶社会的中后期。奴隶主为了有效地对内统治和对外战争,必然要对其掌控范围内的各种资源做到心中有数,由此便产生了对土地、人口及各种资产方面的数据搜集和汇总工作。

中国最早的统计资料是公元前 21 世纪夏朝时期的人口和土地方面的资料。据载,当时中国分为九州,人口约 13 555 万人,土地约 2 438 万顷。在战国时期,中国已经有了“上计”制度,即是按照行政隶属关系逐级上报统计资料的一种制度。而且当时的一些政治家和思想家对这种制度的重要性做过一些论述。秦国的商鞅指出:“强国知十三数,欲强国,不知十三数,地虽利,民虽众,国愈弱至削。”齐国的鲁仲也对此有过精辟地论述,他说:“不明于计数而欲举大事,犹无舟楫而经于水险也。”在国外的一些文明古国,也都有对统计实践活动的记载,无论是古希腊、古罗马,还是古埃及的史料中都可以找到当时人们进行统计工作的证据。如公元前 3050 年,埃及为了建造金字塔就进行了全国范围的人口和财产方面的调查。

统计工作在封建社会时期的发展比较缓慢,其真正的发展是在资本主义社会。由于生产力的极大发展,统治者对国内资源现状的了解欲望更加强烈,而且了解的范围也不仅仅限于过去的人口、土地、税收、粮食等方面,还要了解对工业、农业、商业、银行、保险、交通、邮电、外贸、就业等事关国计民生的方方面面。在欧洲出现了“统计狂热”时期,许多国家建立了统计机关和统计研究机构,很多学者和实践工作者也试图将统计工作的经验进行总结和提升,因此也就有了很多关于统计理论方面的论著。

统计学作为一门学科,大约产生于 17 世纪 70 后代期,以英国人威廉·配第(W. Petty)于 1671—1687 年写成的《政治算术》一书的问世作为标志。该书中已经有了数据计量、表示和显示的方法介绍,虽然比较朴实,但也是今天统计学中的基本内容和方法。威廉·配第以劳动价值理论为基础,对英、法、荷三国的国情国力进行了数量对比分析,对英国的政治和经济发展提出了建设性的意见。马克思对威廉·配第进行了很高的评价,称其为政治经济学之父和统计学的创始人。

在 19 世纪中叶,统计学的发展进入了一个全新的时期,统计学的性质由研究和解决实际问题,即实质性学科向认识论和方法论学科转变。这主要归功于数学尤其是概率论和数理统计的发展和成熟。以比利时人凯特勒(L. Quetelet)为首的一些数理功底非常深厚的统计学者将臻于成熟的概率论原理和方法,如大数定律、中心极限定理、误差法则、正态分布等思想引入到统计学中,从而使这一学科成为了研究大量现象数量方面的一门学科,而不再是为君主、国王等统治者服务的“关于政治事业的数学推理艺术”。

19 世纪以后,统计学有了一个很大的发展,形成了很多分支学科,也出现了不同的学派,但作为一种认识方法论的学科性质并没有改变过。

二、统计学的性质和特点

(一) 统计的含义

统计一词,一般有三种含义:一是统计工作;二是统计资料;三是统计科学。

统计工作是指对社会经济现象的数量方面进行搜集、整理和分析的活动。统计资料是统计工作形成的结果,一般表现为各类数字型资料。统计科学是研究统计过程的理论和方法的科学。上述三种含义并不是孤立的,它们之间存在着密切的联系。统计工作是指人们搜集、整理、分析大量社会经济现象数量表现、数量特征及数量关系的活动,它的产生历史相对久远并且还会继续存续下去;统计工作的成果就是统计资料,没有统计工作也就不会有统计资料,统计资料为各种研究和决策提供必须的数据支持;指导统计工作实践和统计资料形成的理论和方法就是统计科学,它来源于统计工作实践,在统计工作实践中产生和发展,又反过来指导统计工作实践。

(二) 统计学的性质和特点

1. 统计学的性质

统计学作为一门认识方法论学科,以大量现象的数量方面为其研究对象,其阐述的原理、原则和方法,可用来指导统计研究和统计工作,它对统计过程中的核算、分析、组织等环节均有重要的指导作用。统计学作为一门应用性学科,是以大量社会经济现象的数量方面为其研究对象。统计学是通过收集、处理、分析和解释数据从中获得对研究对象进行表述结论的科学。任何一种社会经济现象总会表现为数量和质量两个方面,并且遵循着量变到质变的规律。统计学通过对事物现象量的方面研究来提示其质的方面,探索其发展规律性和必然性。

2. 统计学的特点

统计学的特点主要表现为以下三个方面。

(1) 数量性。统计活动的三个主要环节,即统计调查、统计整理和统计分析,总是围绕着数据的取得、数量关系的确定及分析、变量发展趋势的分析等数量方面进行的,所以统计学的数量性是其最基本的属性,主要表现在以下四点。

① 用横断面的数据表示现象在一定的发展时期所处的规模、结构和水平。根据 2014 年国家统计局所发布的国民经济和社会发展统计公报:初步核算,全年国内生产总值(GDP)为 636 463 亿元。其中,第一产业增加值 58 332 亿元;第二产业增加值 271 392 亿

元;第三产业增加值 306 739 亿元。第一产业增加值占国内生产总值的比重为 9.2%,第二产业增加值比重为 42.6%,第三产业增加值比重为 48.2%。以上关于 GDP 及各次产业增加值就是用于反映规模和水平的数据。而后段文字中关于各次产业增加值占 GDP 的比重即是用于反映产业结构的数据。

② 以时间序列的统计数字反映同一现象总体在不同时间的发展速度和变动趋势。在 2014 年统计公报中也有关于 2010—2014 年全国一般财政收入的数据,见图 1-1。

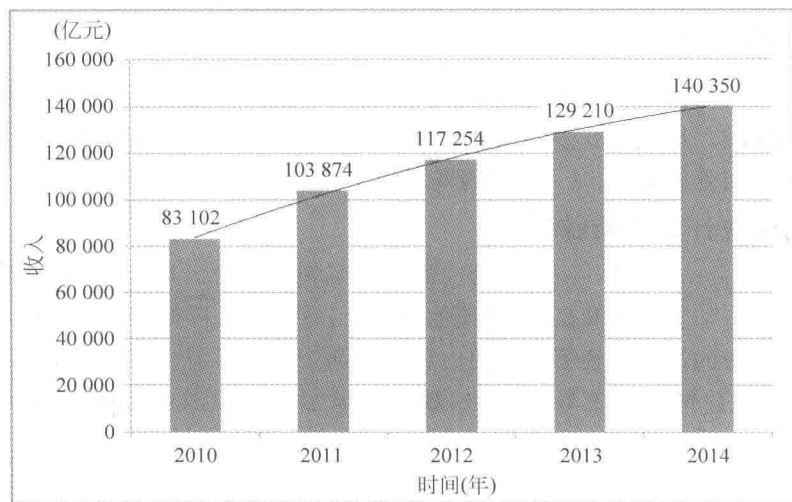


图 1-1 2010—2014 年全国一般财政收入
资料来源:2014 年国民经济与社会发展统计公报。

由图 1-1 可以看出,2010—2014 年中国一般财政收入有明显地增长趋势,但后期的增幅有所减小。

③ 以相关数据的对比反映现象间的联系或存在的问题。例如 1985—2014 年某地区 GDP 平均增长速度为 10.60%,城市职工收入年均增长 6.8%。在这里提到了 GDP 和城市职工收入两个指标的发展变化情况。从数据上看,该地区 GDP 的发展速度较快,而且城市职工的收入也有了相应的增长,这两个指标的同向变化且后者较小,在一定程度上反映了该地区社会经济的良性发展,从数据上看城市职工收入的增长速度低于 GDP 的增长速度也基本符合经济发展的规律。

④ 以历史和现状资料对现象未来的发展规模和水平进行预测。在后面的时间序列分析和相关分析中,将会讨论某一现象的数量随时间变化而变化的问题,也会讨论某一数量的变化受其他变量变化的影响问题。只要找到不同数量变化之间的依存关系,就可以根据影响因素的变化来预测该现象在未来的发展规模和水平了。

(2) 总体性。统计学的研究对象决定了其具有总体性这一显著特点。很多学科的研究是一个从个体到一般的过程,比如医学的研究是从个别病例的研究入手的,管理学的研究也是从个别企业案例入手的。统计学则与这些学科的研究方法有着本质的不同,它关心个别事物现象的质量与数量表现,但它研究的并不是个别事物现象,而是总体的数量特征和数量表现,即是大量现象的规律性和大量过程的倾向性。

(3) 具体性。统计学研究的数量与数学研究的数量有着本质区别, 数学研究的是抽象的数量及数量关系, 而统计学研究的是具有一定社会经济意义的数量。任何事物现象总是表现为质与量的统一, 遵循着量变与质变规律。统计学研究数量首先要对其质量方面有一个准确的认识, 对其质的规定性有了一个正确的认识后, 才能进行数量方面的研究。统计学以数量方面作为研究对象并不意味着统计学的研究目的是数量方面, 其最终的目的是通过对具体数量方面的研究来揭示事物现象质的特征, 探讨其质的规律性。

三、统计学的分科

统计学发展到现在已经形成了许多分支学科, 统计方法也已经被用于自然科学和社会科学的诸多领域。从统计方法的构成来看, 它可以分为描述统计和推断统计; 从统计方法的研究和应用角度来看, 又可分为理论统计和应用统计。

(一) 描述统计和推断统计

描述统计(Descriptive Statistics)研究如何取得客观现象的数量, 并通过图表形式对所获取的数据进行加工处理和显示, 进而通过综合、概括和分析得出反映事物现象规律的数量特征。本书中的第二章到第四章的内容基本属于描述统计学的范畴, 主要研究统计资料的搜集方法、数据的加工处理方法、数据的显示方法以及数据分布特征的分析方法等。

推断统计(Inferential Statistics)则是研究如何从总体中获取一定的样本数据, 并根据从总体中按随机原则抽取的样本数据, 对总体的数量表现和数量特征进行推断的一门科学。一般来讲, 统计推断不是必然性推断, 而是以一定的概率形式给出的推断。

描述统计和推断统计是统计方法论的不同构成部分, 而不是严格意义上的统计学的不同类别。描述统计和推断统计既反映了统计学发展的不同阶段, 也反映了统计工作或统计研究过程的不同阶段。从统计学的发展来看, 早期的统计学基本上属描述统计学, 它更多地关注实际数据的获得和表现, 现代统计学更多地研究数量规律和数量关系, 很多成分属于推断统计。

从统计研究过程来看, 统计研究基本可以划分为在逻辑上互相衔接的三个阶段, 即统计调查、统计整理、统计分析及对数据的解释。统计研究的过程是从统计总体的确定开始的, 经过资料搜集、整理加工环节, 再经过统计分析环节, 完成统计研究的全过程。描述统计和推断统计在这个过程中应用于不同的环节, 描述统计可以说贯穿于统计活动的全过程, 而推断统计则主要应用于统计分析阶段, 它在参数估计与假设检验、相关与回归分析、方差分析、列联分析、时间序列分析和统计指数分析及评价中都可以应用, 见图 1-2。

(二) 理论统计和应用统计

理论统计(Theoretical Statistics)是指统计学的数学原理, 它主要研究统计学的一般理论和一般方法。运用理论统计解决现实问题就形成了应用统计(Applied Statistics)。

理论统计抽象地研究统计的数学原理, 需要广泛的数学基础。由于概率论是统计推断的基础, 所以概率论也可以纳入统计学的范畴。