



21世纪高等院校公共课精品教材



统计学

宋旭光 主编

TONGJIXUE

 东北财经大学出版社
Dongbei University of Finance & Economics Press





21世纪高等院校公共课精品教材

宋旭光 主编

图例目录 (CIP)

2012年出版 东北财经大学 宋旭光 主编 宋旭光 主编

(21世纪高等院校公共课精品教材)

ISBN 978-7-262-42222-9



统计学

宋旭光 主编

TONGJIXUE

1058627

ISBN 978-7-262-42222-9
定价：25.00元

FE 东北财经大学出版社
Dongbei University of Finance & Economics Press

大连

© 宋旭光 2012

图书在版编目 (CIP) 数据

统计学 / 宋旭光主编. —大连 : 东北财经大学出版社, 2012. 9
(21 世纪高等院校公共课精品教材)
ISBN 978-7-5654-0950-9

I. 统… II. 宋… III. 统计学-高等学校-教材 IV. C8

中国版本图书馆 CIP 数据核字 (2012) 第 200774 号

东北财经大学出版社出版

(大连市黑石礁尖山街 217 号 邮政编码 116025)

教学支持: (0411) 84710309

营 销 部: (0411) 84710711

总 编 室: (0411) 84710523

网 址: <http://www.dufep.cn>

读者信箱: dufep@dufe.edu.cn

大连北方博信印刷包装有限公司印刷 东北财经大学出版社发行

幅面尺寸: 185mm×260mm

字数: 411 千字

印张: 17 3/4

2012 年 9 月第 1 版

2012 年 9 月第 1 次印刷

责任编辑: 孙晓梅 石真珍

责任校对: 孙 萍

封面设计: 冀贵收

版式设计: 钟福建

ISBN 978-7-5654-0950-9

定价: 35.00 元

前 言

近年来,我国统计学有了快速的发展,统计作为一种分析工具在经济与社会发展的各个领域有着广泛的应用。可以说,在信息社会中,不管你正在或准备从事什么工作,都要掌握一些统计学的知识,因为当前几乎所有的工作领域都要用到统计方法。

但知易行难,与其他应用性较强的课程比较起来,统计学给人的印象大多是枯燥无味、艰涩难懂,这使得大量有意学习统计学的人望而却步。其实,这是一个误解,学习统计学与学习其他课程一样,都有一个循序渐进、触类旁通的过程,只要静下心来,从基础学起,由浅及深,就一定会在统计学的学习中有非常大的收获。

考虑到这一情况,为了给初学统计学的读者提供一个更加友好的学习界面,有效降低统计学的“入学”门槛,吸引更多的人学习和掌握统计学知识,我们编写了这本《统计学》教材。

本书在编写过程中力求体现以下特点:

一是注重基础。对于初学者而言,打好基础至关重要。只有先入门,才能谈得上登堂入室。因此,本书本着有舍有取的原则,不盲目求大求全,力争在有限的篇幅中尽可能地把统计学基础知识介绍给读者。本书重点是引导读者走进统计学殿堂,使之掌握基本的统计思路,并激发其根据需要进行继续学习的兴趣。为此,我们对本书的内容进行了精选,在不影响整体效果的条件下,弱化了较为复杂和难以理解的内容,强化了统计学基础方法的介绍。

二是形式活泼。统计学来自实践,服务生活,很多统计学知识就在我们身边。本书各章均以“开心一刻”导入,在编写中设置了大量专栏和例题,使读者能够在轻松的环境中更有效率地学习。同时,本书各章还设置了计算机辅助和课堂讨论等环节,使课堂学习与动手实践和互动思考更有效地结合在一起,使读者能更好地理解为什么要学习这些知识,以及如何运用这些知识。

三是模块设计。本书分为10章,各章之间逐步推进但又相对自成体系。第1章主要介绍统计学的内涵和基本概念;第2章介绍统计数据的收集方法;第3章介绍统计数据的整理与显示方法;第4章介绍如何描述统计数据的分布特征;第5章介绍抽样推断和参数估计的基本思想;第6章介绍假设检验和方差分析的基本原理;第7章介绍相关与回归分析的基本内容;第8章介绍时间序列分析的有关方法;第9章介绍统计指数原理及应用;第10章简要介绍经济统计的基础知识。

本书可作为经济学、管理学相关专业的教学与培训用书,也可作为统计工作者和经济管理人员的业务参考书。我们希望有更多的人通过本书喜欢并受益于统计学知识。

本书由宋旭光教授任主编。各章分工如下:宋旭光教授撰写第1章和第10章,刘丹

丹博士撰写第2章和第3章，杨焯军博士撰写第4章，石刚副教授撰写第5章和第6章，吕光明副教授撰写第7章和第8章，陈梦根教授撰写第9章。

在编写本书的过程中，我们参考了国内外部分统计学教材和文献，在此我们对有关作者和文献提供者表示诚挚的谢意。由于时间和水平所限，书中难免有不当或疏漏之处，在此恳请各界读者提出宝贵意见。

编者

2012年8月

目 录

第1章 走近统计学	1
学习目标.....	1
1.1 什么是统计学	1
1.2 统计学起点	5
主要术语	10
课堂讨论	10
课后习题	10
第2章 数据的收集	11
学习目标	11
2.1 数据的含义、作用和分类.....	11
2.2 现存数据的收集.....	15
2.3 原始数据的收集.....	18
2.4 调查问卷的设计.....	28
2.5 量表和量表的设计.....	33
主要术语	38
课堂讨论	39
课后习题	39
第3章 数据的整理与显示	40
学习目标	40
3.1 统计数据整理概述.....	40
3.2 定性数据的整理和显示.....	42
3.3 定量数据的整理和显示.....	48
3.4 绘制统计图表的原则.....	61
主要术语	64
课堂讨论	64
课后习题	64
第4章 数据的分布特征	66
学习目标	66
4.1 数据分布的集中趋势测度.....	66
4.2 数据分布的离散程度测度.....	82

4.3	数据分布的形态测度	89
4.4	利用 Excel 计算描述统计指标	91
	主要术语	94
	课堂讨论	94
	课后习题	95
第 5 章	抽样推断与参数估计	97
	学习目标	97
5.1	几种基本的抽样方法	97
5.2	抽样推断与参数估计中的基本问题	106
5.3	点估计	112
5.4	区间估计	115
	主要术语	122
	课堂讨论	123
	课后习题	123
第 6 章	假设检验与方差分析	125
	学习目标	125
6.1	假设检验基础	125
6.2	总体均值的假设检验	130
6.3	总体比例与总体方差的假设检验	141
6.4	方差分析	149
	主要术语	155
	课堂讨论	155
	课后习题	155
第 7 章	相关与回归分析	157
	学习目标	157
7.1	相关关系的描述与度量	157
7.2	一元线性回归分析	164
7.3	多元线性回归分析	172
7.4	非线性回归分析	177
	主要术语	182
	课堂讨论	182
	课后习题	182
第 8 章	时间序列分析	185
	学习目标	185
8.1	时间序列的动态分析指标	185
8.2	时间序列的成分模型测定分析	193
8.3	时间序列的成分模型预测	205
	主要术语	214
	课堂讨论	214

课后习题·····	215
第9章 统计指数 ·····	217
学习目标·····	217
9.1 统计指数概述·····	217
9.2 综合指数·····	219
9.3 平均指数·····	223
9.4 指数体系·····	228
9.5 常见统计指数·····	231
主要术语·····	242
课堂讨论·····	243
课后习题·····	243
第10章 经济统计基础 ·····	245
学习目标·····	245
10.1 国民经济核算原理·····	245
10.2 国内生产总值统计·····	251
10.3 国民经济统计概览·····	261
主要术语·····	271
课堂讨论·····	272
课后习题·····	272
主要参考文献 ·····	274

第1章 走近统计学

【开心一刻】

三位科学家由法国去意大利参加会议，越过边境不久，发现了一只黑羊。“真有意思”，统计学家谈论道，“意大利的羊都是黑的”。“这种推断并不可靠”，物理学家应道，“我们只能得出这样的结论：在意大利有一些羊是黑色的”。数学家马上接着说：“在意大利，至少有一只羊的至少半侧，在至少一个地方的有些时候，在有些人的眼里，直接用肉眼观察时，看上去是黑色的。”

学习目标

统计学是一个迷人的、广泛应用于各个领域的前沿学科。本章我们将带领大家认识统计学。本章的具体要求包括：①了解统计学的重要性和什么是统计学；②掌握统计学的基本概念；③为进一步学习统计学知识做好准备。

1.1 什么是统计学

现代社会是一个海量的信息社会，越来越多的人用越来越多的时间来面对越来越多的信息，信息的作用越来越重要，如果你面对大量的信息束手无策，那你已经输在起跑线上了。

这些信息就是我们所说的统计数据。用统计学方法正确处理和分析数据，已经成为现代社会通向成功的通行证。正如英国作家威尔斯在一百多年前所预言的那样：“统计终有一天将会与读书写字能力一样重要。”现在，这一天已经来了。

1.1.1 统计学的定义和任务

统计学（statistics）是处理数据的一门科学。具体说来，统计学是研究如何收集、整理、分析、核算数据，并从数据中得出结论的科学。

显然，统计学的主要任务包括：

1. 数据收集

数据收集是统计工作的基础，只有在找到数据之后，才能对数据进行研究。在信息社会，数据无所不在，传感信号、监控录像、商业信息、金融数据等充斥了我们的生活。如何在浩如烟海的数据中找到我们需要的数据，要有专门的方法。本书第2章将介绍这些方法。

2. 数据整理

数据收集上来以后,还要进行必要的整理,比如将数据用图表的形式展示出来,使数据更加直观、清晰和易读。对于大多数人来说,掌握一些必要的数据整理方法是十分有用的。本书第3章将探讨如何进行数据整理。

3. 数据分析

数据分析是指运用统计方法对有关数据进行处理,并从数据中提取出有效信息的过程。如何正确选择和运用有关统计方法对于数据分析至关重要。恰当的数据分析方法可以使人们准确地描述和推断数据所表达的信息,并从数据中得出想要的结论。本书第4章至第8章将介绍这些内容。

4. 数据核算

统计学要在有关学科理论的指导下,通过制定有关定义、标准、方案和体系,为数据的收集、整理和分析提供核算框架。在社会科学领域,为了规范数据的开发与分析,国民经济核算体系、社会核算方法、工商会计等核算技术有着较多的应用。本书第9、10章将讨论有关核算问题。

【专栏1—1】

中国统计学发展的百年历史

我国的统计学教育是从20世纪初的清朝末年开始的,1902年《钦定京师大学堂章程》规定,在商科大学正式讲授统计学课程。

民国期间,随着我国赴欧美留学生人数的增长,越来越多的欧美统计学著作与教材被翻译介绍过来,成为大学中使用的主要教材。当时所教授的统计方法并非只就某一专门学科而言,而是兼收并蓄了经济、商业、教育、心理、生物、生命等各种统计方法。

新中国成立前,我国多数大学中的统计学是作为方法课程在理、工、农、医、商学、社会学、经济学等学科专业开设的。我国最早设立统计系的大学是重庆大学和复旦大学。重庆大学1937年建立了会计统计系,复旦大学1938年在商学院下设了统计学系。

新中国成立后,我国受到苏联的较大影响,将统计学一分为二,即数学中的概率论和数理统计与社会科学中的社会经济统计。在动荡的年代里,统计学受到较大的干扰,发展缓慢乃至停滞。

改革开放后,统计学专业和统计教育得以恢复。1977年,厦门大学、天津财经学院、东北财经学院、山西财经学院和北京经济学院的统计学专业(属经济学类)首批恢复招生。1983年7月,教育部首次批准在华东师范大学、复旦大学、南开大学设立数理统计(属数学类)专业,当年招生。

随着社会对统计学和统计人才需求的不断提升,统计学得到了快速发展。

1992年,国家标准局公布学科分类目录,首次将统计学从经济学中独立出来,成为与经济学、数学并列的一级学科。1998年,当时的国家教委将本科专业目录中的统计学归入理学门类,上升为与经济学、数学等并列的学科。2010年,国务院学位委员会进行研究生专业目录调整,将统计学提升为一级学科,可授理学学位,也可授经济学学位。此外,在我国还有一个“流行病与卫生统计”专业,可授医学学位。

不论是本科生还是研究生的培养,改革开放以来,统计学教育都取得了长足的进步。以中国人民大学本科生招生志愿和录取分数为例,1980—1990年统计学专业录取的第一

志愿学生很少，很多是服从分配的，入学平均分数在各专业中也是中等水平。进入新世纪以来，统计学专业录取的生源质量越来越好，现在基本都是校内第一志愿，录取的平均分数在学校60多个专业中名列前茅5名，特别是二、三年级转专业时统计学已成为学生最愿意选择的专业之一。学生的毕业去向也很好，本科生一半以上出国继续读研或在国内读研。读研和就业的领域都很广，具有广泛的适应性。

资料来源 袁卫：《机遇与挑战——写在统计学成为一级学科之际》，载《统计研究》，2011（11）。

1.1.2 描述统计学与推断统计学

通常可以把统计学分成描述统计学和推断统计学两类。

1. 描述统计学

描述统计学研究如何取得反映客观现象的数据，并通过图表形式对所搜集到的数据进行加工、整理和显示，进而通过综合、概括与分析得出反映事物的数量特征和数量关系的统计方法。

比如，2010年2月10日晚，中国男子足球队在日本东京迎战老对手韩国队，发挥出色的中国队最终以3:0击败韩国队，结束在国际A级赛事中长达32年不胜对手的尴尬纪录，同时也创造了中韩交战史上的最大分差纪录。本场的技术统计如表1—1所示。

表1—1 2010年东亚男子足球四强赛技术统计表（中国—韩国）

中 国	数 据	韩 国
3	入 球	0
6	射 门	23
4	射 正	7
17	犯 规	15
37.4%	控球率	62.6%
3	越 位	1

通过表1—1中的描述性统计数据，我们可以对比赛的情况有一个数量方面的认识。

2. 推断统计学

推断统计学是研究如何利用样本数据来推断总体特征的统计方法。它是在对搜集的样本数据进行描述的基础上，对统计总体的未知数量特征作出以概率形式表述的推断。其中，总体是指我们所要研究事物的全体，样本是指从总体中抽取的一部分单位的集合（详见下一节）。

比如，要了解消费者对某种食品的口味偏好，我们不可能对每个消费者一一进行测试，而只能选取一部分消费者，通过这部分消费者的情况对全体消费者的情况进行推断。这就用到了推断统计学。

描述统计学与推断统计学的划分，反映了数据分析的不同层次。如果我们所处理的数据是一个总体的数据，则可以通过描述统计学认识该总体的数量规律性；如果我们所处理的数据只是总体的一部分，那么就必须通过推断统计学对总体的数量规律性进行科学的推断。

描述统计学与推断统计学都是统计学的重要组成部分，描述统计学是统计学的基础，推断统计学是统计学的主干。随着数据量的不断增多和统计任务的不断复杂化，统计工作的重心已经从对过去的回顾转移到对现状的分析和对未来的预测上来。在当前情况下，大量的统计工作所搜集的数据是样本数据，因此，推断统计学在统计学中的地位和作用日益增强。当然，这并不意味着描述统计学已经不重要了，事实上，描述统计学在日常工作和生活中一直有大量的应用，而且如果没有描述统计提供有效的样本信息，推断统计也难以进行。

1.1.3 统计学的应用领域

统计学有着良好的发展前景，只要有数据的地方就会用到统计方法。

可以说，在信息社会，几乎所有的工作领域都要用到统计方法。比如，在医学界，人们利用统计方法来研究疾病的原因或影响因素、判断药物和治疗方法的有效性；在考古界，人们利用统计方法来推断特定文物的历史年代；在心理学界，人们利用统计方法来分析特定刺激的心理效应；在气象学界，人们利用统计方法来预测未来的天气情况；在物理学界，人们利用统计方法对由大量粒子组成的宏观物体的物理性质及宏观规律作出微观解释；在生物学界，人们利用统计方法来研究基因定律……

具体到经济与管理领域，统计学的应用也是十分广阔的，下面简要介绍其中的一些。

1. 公共管理领域

决策科学化使统计学在公共管理部门的应用不断增强。对于公务人员来说，如果不能掌握一些数据搜集和分析技能，将很难适应政府管理信息化、科学化的需要。

2. 经济研究领域

经济研究正在走向定量化和数理化。在进行经济分析和预测时，必然要用到各种统计信息和统计方法。

3. 产品质量管理

质量是企业的生命，质量管理与控制已经成为统计学在生产领域的一项重要应用，各种统计方法被广泛应用于生产过程的监测。

4. 市场调研领域

企业要及时应对市场的变化，以在竞争中取得优势。而了解市场必然要涉及市场信息的搜集和分析过程，在这方面，统计方法的应用极为广泛。

5. 人力资源管理

人们可借助统计方法全面掌握本单位的人力资源状况，设计更为合理的绩效考评机制。

6. 金融保险领域

随着金融保险行业的快速发展，金融与保险数据空前膨胀。相关数据处理工作使金融保险行业对统计人才的需求与日俱增，保险精算人才更是供不应求。

【专栏1—2】

当今大学生的关键词：统计学

Carrie Grimes 在哈佛大学主修人类学和考古学专业，她去过很多地方。在洪都拉斯，她曾通过绘制古文物发现地的地图来研究玛雅人的居住模式。“人们认为野外考古学就像印地安娜·琼斯（影片《夺宝奇兵》主角）的冒险活动一样，可实际上，我们的很多工

作都是在进行数据分析。”她表示。

Grimes 目前就职于谷歌 (Google) 公司, 她现在是谷歌的统计分析师, 每天要面对大量的数据, 并运用各种统计分析方法寻找能改善公司搜索引擎效果的方法。

一般人对网络时代的印象大多是成天与电脑为伍、足不出户的宅男宅女, 而 Grimes 女士是网络时代的统计学家, 她和其他同属网络时代的优秀同伴们一样, 正在致力于扭转大众对网络时代的刻板印象。在她们的影响下, 越来越多的人认识到统计工作不仅被社会迫切需要, 而且统计人才也变得越来越抢手。

“我一直说在未来十年里, 最迷人的工作将是统计工作, 这可不是开玩笑。”谷歌的首席经济学家 Hal Varian 表示。

IBM 公司在数据搜索服务中发现了商机, 它创建了商业分析和优化服务小组。该小组将在调查实验室中招收 200 名数学家、统计学家和其他数据分析师——相对于 IBM 计划在整个公司中培训或新招超过 4 000 名分析师来说, 200 这个数字并不算多。

从其他方面也能看出人们对统计学的兴趣正在不断增强。据美国统计协会表示, 2009 年在华盛顿举行的统计学年会的出席人数将达 6 400 人, 超过往年的 5 400 人。与会的人们看似与在华盛顿的一群群游人毫无区别, 但仔细观察你会发现他们正全神贯注地讨论的话题是随机性、参数、回归和聚类等统计学术语。

数据的爆发式增长也提升了一些职业的地位, 保险公司的精算人员就是一个明显的例子。

资料来源 节选自 Steve Lohr, For Today's Graduate, Just One Word: Statistics, New York Times. August 6, 2009。

可想而知, 作为一种通用的数据处理方法, 统计学已经成为一种生活和工作中的必备工具。越来越多的人认识到, 必须学习和掌握一定的统计学知识, 只有这样, 才能有效应对信息时代和网络经济带给我们的挑战。

1.2 统计学起点

通过上面的介绍, 大家可以对统计学是什么, 以及统计学有什么用有一个大致的印象, 这是一个起点。千里之行, 始于足下, 让我们开始学习统计学的知识吧!

1.2.1 统计学的基本概念

统计学要用到许多概念, 有一些基本概念是经常用到的, 这里先加以介绍。

1. 确定性与随机性

我们在学习数学、物理、化学等课程的时候, 大多数情况下是在学习确定性的规律。也就是说, 在一定条件下, A 必定会导致 B 的产生。比如, 在一定约束条件下, 该方程有唯一解; 在标准大气压下, 水的沸点是 100°C ; 在一定条件下, 化学反应结果必然产生。事实上, 我们也运用这些确定性的规律解决了很多科学问题。

但确定性不是全部, 生活中有大量事情不是必然发生的, 无法用确定性的理论去分析它。比如, 每天乘坐公交车的人数、股票价格走势、某种疾病的发病率等。最典型的例子是抛硬币, 我们无法确切地猜出下一次是什么结果。这些事例都受到不确定的随机因素的影响, 具有随机性。

随机性并不意味着混乱，事实上，许多随机事件是有规律的。比如，每天早高峰出现的时间是大致相同的，股票价格指数从长期来看与宏观经济趋势基本一致，某种疾病的发病率有明确的规律等。抛硬币更是这样，随着次数的增加，正反面出现的概率是相等的。

可以说，随机事件尽管受不确定性因素影响，但也有一定的规律性，通过科学方法找到随机变量变动的规律性并据此作出分析和预测，是统计学研究的重要使命。

2. 总体与样本

所谓总体 (population)，是根据一定目的确定的所要研究的事物的全体。它是由客观存在的、具有某种共同性质的许多个别事物构成的整体。组成总体的每个个别事物称为个体，也称为总体单位。总体中个体的数量被称为总体容量。

总体具有大量性、同质性和差异性三个特征。大量性是指总体中的个体数量要足够多；同质性是指总体中的各个个体必须具有某种共同属性或特征；差异性是指构成总体的各个个体在某些方面又必须存在一定的差异。

总体根据其包含的个体数量是否可数可以分为有限总体和无限总体。比如，某大学的学生数量是可数的，其学生总体就是有限总体。而对空气质量的研究来说，空气就是无限总体。

统计学要研究总体的特征，但是，当总体中的个体数量很多，甚至无限时，就没有必要或无法对构成总体的所有单位进行研究了。这时，可以抽取总体中的一部分个体，作为总体的代表加以研究。这种由总体的部分个体组成的集合被称为样本 (sample)。样本中个体 (样本单位) 的数量被称为样本容量。

3. 参数与统计量

用来描述总体特征的某种特征值被称为总体参数 (parameter)，在统计学中，总体参数通常用希腊字母表示。

用来描述样本特征的某种特征值被称为样本统计量 (statistic)，在统计学中，样本统计量一般用英文字母表示。

总体参数往往是很难获取的，而样本统计量是可知的，因此，用样本统计量去估计总体参数是统计学研究的一项重要工作。

【专栏 1—3】

统计使人豁达

哲人培根曾经说过：“历史使人明智，诗歌使人巧慧，数学使人精细，博物使人深沉，伦理使人庄重，逻辑与修辞使人善辩。”按照这个套路下来，我想加上一句，这就是：统计使人豁达。

那么具体说来，统计究竟对人生有什么好处呢？

我想摆在第一位的应该是，良好的统计训练使人宏观意识比较强。谁都知道，总体和总量是统计学中最基本的概念，虽然统计调查和处理都要从个体入手，但其最终目的是由此去对现象的总体作出评价。统计总是把着眼点放在对宏观大势的把握上，对个体数据如何处理，也取决于怎样科学地得到对总体的认识。久而久之，统计训练使人宏观意识得到强化，遇到问题习惯将之放在时空的大背景中去考虑。比方说，判断学校某个学科的地位，不仅要将其与校内其他学科相比较，更要将其与国内其他院校的相同学科进行比较；

在制定长期发展目标时，还要与发达国家的该学科水平相比较。只有在大系统中去把握，所得出的结论才会是可靠的。先人云：会当凌绝顶，一览众山小。统计训练给人的就是这种感觉。

良好的统计训练还能使人在看待和处理问题时不那么偏激。这是不难理解的，统计中的平均分析和平衡分析都在教导我们对事物要持中庸的态度，搞统计的人的豁达也缘于这一点。知道了平均水平，便把握了事物的一般趋势，心中就有底。同时也知道事物围绕平均水平还会有变异，当好的极端出现时，能受宠不惊；而当差的极端出现时，又可视之泰然，胸襟于是开阔通畅。

上面说的是统计训练对人的时空方位感的作用，有了方位感，便容易豁达。下面想说的是统计训练使人的豁达有较为坚实的基础，这就是对人的分寸感的培养。统计注意研究事物量变引起质变的度，注意研究事物的数量规律。统计是用数据说话的，不能捕风捉影，得出的结论通常有方法依据和信息支持，比起单纯的理论说教，更能给人以实在的感受，分寸感强。

总而言之，现代社会对统计是很认可的。君不见，诺贝尔经济学奖得主有多少本身就是统计学家，又有多少在其研究中大量运用或依靠了统计！进入新世纪以来，我国统计学发展也进入了快车道，统计的专业优势开始得到社会的广泛认可。

资料来源 邱东：《统计使人豁达》，载《偏，得以见》，桂林，广西师范大学出版社，2010。

4. 变量与指标

(1) 变量。

变量 (variable) 是统计学中的重要概念，统计研究往往从研究对象的某些特征开始，比如，要了解一支球队的情况，我们可以从队员的籍贯、年龄、身高、体重、进球数量等特征入手。显然，不同队员的上述特征是不尽相同的，我们称之为变量。

准确地说，变量是说明现象某一特征的概念。变量的具体取值就是变量值，比如一个球员的身高是180cm，这里的180cm就是“身高”这一变量的“变量值”。

根据变量值是否连续出现，变量可分为连续型变量和离散型变量。连续型变量是指变量的取值在数轴上连续不断，无法一一列举，比如上述“身高”、“体重”等变量。离散型变量是指变量的数值在数轴上是不连续的，只是一个个独立的点，比如上述的“进球数量”。

受随机因素影响的变量被称为随机变量。比如，球员每场进球数要受到竞技状态、裁判、天气、队友配合等多方面不确定因素的影响，它是一个随机变量。再比如，公共汽车站的候车人数也是随机变量。

在某些场合，变量有时也被称为标志，变量值也常被称为标志值。

(2) 指标。

指标是反映总体数量特征的概念和数值。比如，球队的平均身高是185cm，其中，“平均身高”是概念，“185cm”是数值，既有概念又有数值的“平均身高185cm”是指标。

在日常生活中，经常可以见到各种各样的统计指标。比如，“2011年，我国国内生产总值471 564亿元，比上年增长9.2%。其中，第一产业增加值47 712亿元，增长4.5%；

第二产业增加值 220 592 亿元, 增长 10.6%; 第三产业增加值 203 260 亿元, 增长 8.9%。第一产业增加值占国内生产总值的比重为 10.1%, 第二产业增加值比重为 46.8%, 第三产业增加值比重为 43.1%”。

从结构上看, 指标包括变量和变量值, 但要注意, 指标更加强调对总体的描述, 而变量和变量值往往是针对个体而言的。

客观现象是错综复杂的, 为了更全面地反映总体各方面的情况, 有时需要通过多个不同指标来研究总体。这些由反映同一总体多个方面数量特征的、一系列相互联系的指标所组成的体系, 被称为指标体系。

1.2.2 统计软件

处理数据不是一件容易的事情, 在我们运用复杂统计方法处理海量数据的时候更是这样。

好在我们生活在计算机时代, 原本一些枯燥、繁琐的数据处理工作可以借助计算机完成了, 这真是统计学学习者的一个福音。在某些场合, 我们只需要知道统计学的原理, 选择合适的统计学方法, 就可以处理非常复杂的统计分析任务了。因此, 运用统计软件非常重要。

统计软件有很多, 我们介绍常见的一些。

1. Excel

Excel 是微软办公套装软件的一个重要组成部分, 它可以进行各种数据的处理、统计分析和辅助决策操作, 广泛应用于管理、统计、财经、金融等众多领域。Excel 可以胜任基本的数据整理和统计图表的制作工作, 能够实现一些基本的统计分析功能。如果你不从事统计领域的专业工作, 那么 Excel 是你理想的选择。由于 Excel 普及性比较高, 我们在本书中涉及数据处理的软件实例时, 基本上以 Excel 2007 为例。

2. SPSS

如果你觉得 Excel 还不够酷, 但又不愿意在统计软件上花太多的时间, SPSS 比较适合你。SPSS 的全称是“统计产品与服务解决方案”(Statistical Product and Service Solutions), 它是世界上最早采用图形菜单驱动界面的统计软件, 最突出的特点就是操作界面极为友好, 输出结果美观漂亮。它将几乎所有的功能都以统一、规范的界面展现出来, 使用 Windows 的窗口方式展示各种管理和分析数据方法的功能, 对话框展示各种功能选择项。用户只要掌握一定的 Windows 操作技能, 粗通统计分析原理, 就可以使用该软件从事统计分析工作。2009 年 7 月 28 日, IBM 以 12 亿美元现金收购了 SPSS, 如今 SPSS 的版本已出至 20.0, 而且更名为 IBM SPSS。

3. SAS

SAS (Statistical Analysis System) 是由美国北卡罗来纳州立大学 1966 年开发的统计分析软件, 经过多年的完善和发展, 在国际上被誉为统计分析的标准软件, 在各个领域得到广泛应用。全球 50 000 多家企业都在通过 SAS 软件对数据进行深入挖掘, 帮助企业更快、更准确地进行业务决策。与 SPSS 比较起来, SAS 更加专业, 需要使用者对所使用的统计方法有清楚的认识, 并允许使用者通过输入程序实现更为准确的统计分析。对于非专业人士来说, SAS 有一点点难。

【专栏1—4】

其他常用统计软件

除了 SPSS 和 SAS 等统计软件外，S-Plus、R、Stata 和 Eviews 也很受欢迎。

S-Plus 是由美国 MathSoft 公司开发的一种基于 S 语言的统计软件，主要用于数据挖掘、统计分析和统计作图等。S-Plus 的最大特点在于它可以交互地从各方面发现数据中的信息，并可以很容易地使用一种新的统计方法。另外，S-Plus 的数据可以直接来源于 Excel、SAS、SPSS 等软件，其兼容性极好。

R 是一个免费的、由志愿者管理的统计分析软件。它几乎是 S-Plus 的一个克隆（不要钱的 S-Plus），几乎所有从 R 中学到的都可以在 S-Plus 中应用，反之亦然。R 是一个自由软件，在 R 主页（<http://www.r-project.org>）那儿可以下载 R 的安装程序、各种外挂程序和文档。

Stata 是一套提供数据分析、数据管理以及绘制专业图表的完整及整合性统计软件。Stata 可以通过网络实时更新每天的最新功能，更可以得知世界各地的使用者对 Stata 提出的问题与解决之道。Stata 的统计功能很强，具有很强的程序语言功能，用户可以充分发挥自己的聪明才智，熟练应用各种技巧。

Eviews 是 Econometrics Views 的缩写，直译为计量经济学观察，通常称为计量经济学软件包。它的本意是对社会经济关系与经济活动的数量规律，采用计量经济学方法与技术进行“观察”。Eviews 是美国 QMS 公司研制的在 Windows 下专门从事数据分析、回归分析和预测的工具。使用 Eviews 可以迅速地找出数据中的统计关系，并用得到的关系预测数据的未来值。Eviews 的应用范围包括科学实验数据分析与评估、金融分析、宏观经济预测、仿真、销售预测和成本分析等。

目前，常用的统计软件都有专门的教材供我们学习使用方法，大家可以在学习中参考。

1.2.3 怎样学好统计学

你可能觉得统计学听起来很难、学起来压力很大，其实情况并不是这样的，在进入正式的课程之前，我们建议大家放松心情、快乐学习。

(1) 统计学很重要，要认识到统计学不只是一门课程，学习统计学不是为了应付期末考试。学习统计学的根本目的是跟上时代前进的步伐，提升自己的竞争力。可能有一些课程学过了就过去了，以后根本用不到，但统计学不会。

(2) 循序渐进。现在小学一年级的数学课本中就有统计学的学习单元了，从小学生成长为统计学家是一个循序渐进的过程，不可能通过一门课程就完全掌握统计学的知识。事实上，统计学也是有不同层次的，我们可以不选择成为统计学家，但统计学的知识能多学一些就多学一些。

(3) 不要怕，我们不是为了让你记住那些复杂的定义和变态的公式，而是为了让你懂得那些公式后面的精彩的统计学思想。而且，这些思想不只是在统计中运用，你要知道，其实，很多学问都是相通的，你会弄清楚的。

(4) 坚持才会胜利。有些时候，必须下定决心来成就一件事情。偶尔偷个懒是可以的，但不要让惰性主宰你的生活。要记住，跟随老师学习是最快捷有效的方法，要充分利用上课的时间，不要逃课。