



普通高等教育“十一五”国家级规划教材



教育部经济管理类核心课程教材

# Statistics

# 统计学

## ——思想、方法与应用

► 袁 卫 刘 超 编著

普通高等教育“十一五”国家级规划教材

教育部经济管理

# Statistics

# 统计学

——思想、方法与应用

►►袁 卫 刘 超

中国人民大学出版社  
· 北京 ·

**图书在版编目 (CIP) 数据**

统计学：思想、方法与应用/袁卫，刘超编著. —北京：中国人民大学出版社，2011.9  
普通高等教育“十一五”国家级规划教材. 教育部经济管理类核心课程教材  
ISBN 978-7-300-14345-3

I . ①统… II . ①袁…②刘… 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 (2011) 第 185861 号

**普通高等教育“十一五”国家级规划教材**

**教育部经济管理类核心课程教材**

**统计学——思想、方法与应用**

袁 卫 刘 超 编著

Tongjixue: Sixiang, Fangfa yu Yingyong

---

**出版发行** 中国人民大学出版社

**社 址** 北京中关村大街 31 号

**邮 政 编 码** 100080

**电 话** 010-62511242 (总编室)

010-62511398 (质管部)

010-82501766 (邮购部)

010-62514148 (门市部)

010-62515195 (发行公司)

010-62515275 (盗版举报)

**网 址** <http://www.crup.com.cn>

<http://www.ttrnet.com> (人大教研网)

**经 销** 新华书店

**印 刷** 北京华正印刷有限公司

**规 格** 185 mm×260 mm 16 开本

**版 次** 2011 年 9 月第 1 版

**印 张** 23.75 插页 1

**印 次** 2011 年 9 月第 1 次印刷

**字 数** 508 000

**定 价** 39.80 元

---

# 教育部经济管理类核心课程教材

## 出版说明

按照购买力平价标准衡量，中国已被世界银行列为世界第二大经济体，仅次于美国。但是，我们不能因此沾沾自喜。成为经济大国并不意味着就是经济强国，中国的强国之路依然漫长而曲折。我们应该清醒地认识到，面对新的发展形势，我们自身还存在着许多短板，如果不能及时将这些短板补齐，我们将会在前进的道路上失去平衡并摔跤。最重要的短板之一，是我们在经济管理高等教育与实践方面的落后和不足。中国现代经济管理实践比西方国家晚几十年甚至上百年，很多理论知识和实践经验最初是从西方“拿来”的，这导致中国的经济管理类人才在知识储备上总是落后于人，缺乏领先的理念来引导实践。

基于以上认识，中国人民大学出版社近年来不断深化教材的层次和结构，无论是引进版还是本版，都从多个维度进行开发和建设，以适应新的发展要求。作为国内最早引进国外优秀经济管理类图书的出版社之一，我们最初引进的一批经典欧美经济管理类图书造就了一大批成功的管理者。借鉴引进版的成功经验，在本土教材开发方面，除了及时吸纳国内外经济管理领域的先进思想和理念，还提供尽可能多的案例，特别是本土案例。这一点在“教育部经济管理类核心课程教材”系列中体现得十分充分。

本套教材的开发思路得到了全国许多经济管理类高等院校的优秀老师的极大认同和支持。感谢这些老师投入极大的热情，与我们共同设计整套教材的方案，制定教材开发原则和体例。并积极承担各自领域教材的编写工作。每位参编老师都是各自领域的佼佼者，并且无论其身居何职，都依然站在教学第一线。我们尽力做到教材从内容到形式都具有独特的风格；同时，我们还为许多教材配备了案例集或学习指导书，并提供一些教学辅助资料供老师免费下载，为使用教材的老师和学生们提供尽可能周到的服务。

作为新中国成立后最早建立的一家大学出版社，中国人民大学出版社一直秉承“出教材学术精品，育人文社科英才”的宗旨。如今同类经济管理类教材充斥市场，我们更觉得有责任紧跟时代脉搏，不断推出精品，提升教材的质量和层次，一方面，为选择教材的广大师生节约选书的时间成本，另一方面，也希望为提升中国的经济管理教育和实践水平做出贡献。我们期待着广大使用者的建议和鞭策，促使我们不断对本套教材进行改进和完善，使之长远传承，经久不衰。

中国人民大学出版社

# 前 言

为了满足当代学生和广大读者对统计方法在信息社会中日益增长而又尚未满足的需求，我们专门为各领域应用统计和数据分析编写了这本教材。

随着统计知识的普及以及经济、社会的发展，特别是在信息化社会的今天，理、工、农、医、经济、管理以及人文社会科学对统计数据的需求在持续增加，需要处理大量的数据信息，统计是有力的分析工具和决策方法。

大多数统计应用不是由统计学家实施的，而是实际工作者完成的，教师首先应该了解学生的想法和需要，才能达到教学相长。但是很多统计教师在对学生或者实际工作者讲授统计时，总是习惯性地希望教学对象也运用自己学习统计时的思维方式，这是一个误区！

目前我国统计教学过程中存在两个困难或误区：(1) 国内的很多教材过于重视数学公式和推导，使得读者轻视对统计思维的培养，导致面对实际数据时感到恐惧而束手无策；(2) 现有很多教材虽然也强调应用，但是忽略统计方法的软件使用，虽然可以通过软件操作手册掌握软件的使用，但这增加了读者学习的负担，而且这些使用手册过于强调操作过程，忽略对软件输出的统计结果的解释，导致读者不能正确解读统计分析结果，面对一大堆输出而束手无策。

本书作为一本统计学入门教程，体现了以下准则：

- 读者通过分析实际数据，解决实际问题，得到最大的启发。
- 读者应该有机会去使用统计软件分析真实的数据。
- 相对于概率论的基础知识的学习，我们更应该重视概率意识的建立。
- 强调统计的应用，而非统计的理论知识。

本书的最大特色就是应用性强，不仅介绍各种统计方法并且注重在实际中的应用。以大量实际数据运用为基础，为读者深入领会统计思想与方法提供了最有效的工具，编排新颖，讲解清晰。每一种统计方法的介绍都有许多例子和案例，并在每章配有大量的例题和练习。这些统计信息不仅能使读者产生兴趣，还可以让读者从

中学习统计方法及其应用。

本书的另一特色就是可读性强，语言轻松活泼，通俗易懂。本书借助于界面越来越方便友好的统计软件，避免了烦琐的数学推导和统计公式，采用深入浅出、循序渐进的语言系统介绍统计学的知识，帮助读者从中得到有用的信息，并且能够很容易地利用 SPSS 和 Excel 完成各种统计分析的运算。其间穿插了大量丰富、鲜活、生动的扩展阅读，使得统计技术不再枯燥，方便了读者的理解和学习。

本书的特色具体体现在以下几个方面：

1. 案例引子。所有的新知识都通过实际数据的例子来阐述。各章安排了“案例引子”，通过实际生活中的各种实例引导读者进行学习。既可以激发读者学习的兴趣，又可以将复杂的事物简单化、具体化，帮助读者更好地理解和学习统计方法与知识，让读者逐步掌握解决实际问题的能力。

2. 学习目标。在每章的“学习目标”部分，指明教学时应该明确的相关问题。帮助读者系统、全面、深入地理解和掌握基本概念、理论和方法。

3. 想一想。每章均安排了“想一想”的小习题，其中的概念和问题涉及的方式都与示例相似，通过这部分练习，读者可以实时测验自己对知识的理解，巩固所学知识。如果你肯花上一点儿时间把书中讨论的话题和你自己的实际经验结合起来，将会更清楚地认识到，你对统计理解了多少，哪里还存在问题，该如何把课堂内的体验和课堂外的生活联系起来，等等。这种行为将引起你对统计更持久的学习。

4. 专栏或扩展阅读。每章均安排了“专栏”或“扩展阅读”，给出一些与统计知识有关的历史故事或注意事项，帮助读者了解所学知识的重要性和实践性。

5. 无公式化教学。计算机和教学的结合，尤其是界面越来越方便友好的统计软件的使用，已经使旧的学习方法——特别是记忆并运用统计公式——不再适用于大部分读者。本书在每章的讲解中尽量不使用统计公式，而是把统计公式放在每章末的附录里。

6. 历史上的统计学家简介（人物小传）。本书简要介绍了一些著名的统计学家以及他们的成就。通过对他们的了解，扩展读者的视野，使读者了解统计学发展历史上重要的人物以及学科发展历史事件，读者会更加全面地了解统计这个学科。

7. 真实的应用案例。每章后面都有应用案例，学生需要分析案例中的数据去解决所提出的问题，方便学生利用本章所学知识解决实际问题。

8. 挑战性的案例分析。每两到三章统计知识结束之后安排一个挑战性的案例分析，读者需要用前面章节中提出的方法去分析这些案例，复习学过的知识。使得读者借此拓展自己的思维，挑战自己运用统计的能力。

9. 用统计软件分析数据。书中所有的统计分析方法，都运用了使用最为广泛的两种统计软件（Excel 和 SPSS）的分析输出结果来进行讲解。读者在这个高科技的商业社会中肯定要接触这些统计软件输出结果。

10. 统计软件指南。在每一章的最后都给出了“软件的使用”，介绍如何使用 SPSS 和 Excel 操作该章学习到的统计分析方法。读者可以从中得到有用的信息，

并且能够很容易地利用 SPSS 和 Excel 完成各种统计分析的运算。

本书可作为理、工、农、医、经济、管理、人文社会科学专业以及其他领域的统计学教材，也可以供从事商务活动和经济分析等实际工作的各类人员参考。读者可以根据时间和需要，有选择地学习有关内容。本书所采用的数据和课件可以在中国人民大学出版社工商管理出版分社网站（[www.rdjg.com.cn](http://www.rdjg.com.cn)）下载。

由于本书是作者第一次编写通俗易懂的统计教材，疏漏和错误在所难免，恳请读者批评指正，以便再版时修改完善。希望本书能引起读者对统计的兴趣，不再惧怕统计，并帮助读者读懂数据。

袁卫 刘超

2011 年 8 月

# 目 录

<b>第 1 章 统计学导论 .....</b>	1
1. 1 什么是统计学 .....	4
1. 2 如何运用统计学 .....	5
1. 3 统计学的基本概念 .....	7
1. 4 数据的收集 .....	10
1. 5 数据阵/数据文件 .....	17
1. 6 统计学和数学、统计软件的关系 .....	18
附录：软件的使用 .....	19
练习 .....	28
应用案例 .....	29
挑战性的案例分析 .....	31
<b>第 2 章 分类数据的描述分析 .....</b>	33
2. 1 数据分析的三个原则 .....	35
2. 2 频数表 .....	35
2. 3 统计图 .....	37
2. 4 列联表 .....	46
附录：软件的使用 .....	54
练习 .....	56
应用案例 .....	59
挑战性的案例分析 .....	59
<b>第 3 章 定量数据的描述方法 .....</b>	61
3. 1 展示数据的分布 .....	62
3. 2 定量数据的其他图形表示 .....	71



3.3 分布的集中位置 .....	77
3.4 分布的离散程度 .....	86
3.5 分布形态的度量 .....	91
3.6 标准得分 .....	93
3.7 利用描述方法扭曲事实 .....	94
附录：软件的使用 .....	96
公式 .....	97
练习 .....	98
统计图应用的经典案例 .....	103
应用案例 .....	106
挑战性的案例分析 .....	107
<b>第4章 概率 .....</b>	<b>111</b>
4.1 什么是概率 .....	112
4.2 怎样求概率 .....	113
4.3 概率的计算 .....	116
4.4 离散随机变量的分布 .....	118
4.5 连续随机变量的分布 .....	125
4.6 使用概率来检验假设 .....	133
附录：软件的使用 .....	135
公式 .....	140
练习 .....	142
应用案例 .....	145
<b>第5章 参数估计 .....</b>	<b>147</b>
5.1 样本统计量和总体参数 .....	149
5.2 点估计 .....	150
5.3 抽样分布与中心极限定理 .....	153
5.4 区间估计：给结论留一些余地 .....	158
5.5 合适样本量的确定 .....	173
附录：软件的使用 .....	175
公式 .....	176
练习 .....	180
应用案例 .....	183
挑战性的案例分析 .....	184
<b>第6章 假设检验 .....</b>	<b>187</b>
6.1 作为一个问题的假设 .....	188



6.2 怎样回答零假设所提出的问题 .....	191
6.3 显著性水平 .....	195
6.4 正态总体均值的假设检验 .....	198
6.5 总体比例的假设检验 .....	206
6.6 假设检验和置信区间 .....	209
6.7 何时拒绝零假设 .....	210
附录：软件的使用 .....	212
公式 .....	214
练习 .....	215
应用案例 .....	218
挑战性的案例分析 .....	220
<b>第 7 章 方差分析 .....</b>	<b>221</b>
7.1 单因素方差分析 .....	223
7.2 方差分析回顾 .....	236
7.3 双因素方差分析 .....	236
附录：软件的使用 .....	241
公式 .....	243
练习 .....	244
应用案例 .....	248
<b>第 8 章 相关与回归分析 .....</b>	<b>250</b>
8.1 两个变量间是什么关系 .....	252
8.2 两个变量间的关系强度 .....	255
8.3 回归分析 .....	257
8.4 总体中的关系 .....	266
8.5 多元回归分析 .....	268
8.6 虚拟变量 .....	277
附录：软件的使用 .....	280
公式 .....	281
练习 .....	283
应用案例 .....	291
挑战性的案例分析 .....	293
<b>第 9 章 时间序列与指数 .....</b>	<b>296</b>
9.1 时间序列的成分 .....	297
9.2 平滑技术 .....	300
9.3 趋势分析 .....	306



9.4 指数 .....	311
附录：软件的使用 .....	324
公式 .....	326
练习 .....	327
应用案例 .....	330
<b>第 10 章 国民经济统计初步 .....</b>	<b>333</b>
10.1 概述 .....	334
10.2 国内生产总值（GDP） .....	335
10.3 国民收入指标 .....	345
10.4 价格指数和总体价格水平 .....	347
10.5 失业率 .....	349
10.6 国际收支 .....	351
练习 .....	352
应用案例 .....	354
统计术语 .....	356
附录 1 标准正态分布表 .....	362
附录 2 $t$ 分布表 .....	364
参考文献 .....	365

# C 第1章

## Chapter 1 统计学导论

### 一生一世的统计数字！

《今日文摘》2009年第1期有篇文章叫“一生一世的统计数字！”文中有这样一段话：“一个人赤条条来到世界，离开的时候，也带不走多少东西。但是，一个人的一生给地球留下了什么？又创造了什么？以下是人活一世的一组数据。数据以英国人的生活方式为标准，兼顾了世界各地的人。这组数据可以给我们一个参照，也可以给我们很多思考。英国人平均寿命78.5岁，共24亿零750万秒……人一生吃掉的东西：4头牛，15头猪，21只羊，1200只鸡，13000枚鸡蛋（未出生的鸡），5000多只苹果，10000多个胡萝卜，3吨面包，630公斤巧克力，2吨葡萄酒，11吨啤酒（全球随时都有4500万个醉鬼），18吨牛奶，75000杯茶，相当于装满一个浴缸的罐头豆子，一生总共吃下约50吨食物。当然，这是指世界各地人们的平均数。考虑到很多穷人没什么吃的，富人应该吃的更多……一生认识的人（有两年以上的交往）约1700人，长期社交圈约300人……每天说4300个字，一生大约说1.2亿个字词，大多没什么意义……一生读报纸1.5吨；一生读500本书，考虑到有40%的人从来不看书，爱读书的人一生读书超过1000本。一个人一生读的书和报纸，至少需要24棵树。你种了几棵？一生2900多天在看电视（按24小时算），差不多在电视机前不睡觉不说话坐了8年……一生做梦10万次，还不包括白日梦。”

这些数据未必都准确，但信息很清晰。这里既有无法改变的，也有可以改变的；既有真实的生活，也有背后的贫富差距。

我们如何看待这些统计数据，其实就是如何看待自己，也是如何看待人类。是增加这些数据，还是减少这些数据，不光对我们自己，对人类都会产生影响，这就是统计科学的力量。



## ... 学习目标

- (1) 用一句话说明统计学；
- (2) 理解我们为什么要学习统计学；
- (3) 懂得如何运用统计学；
- (4) 知道统计学的基本概念；
- (5) 懂得什么是描述统计学和推断统计学；
- (6) 了解数据如何收集以及数据收集中存在的问题；
- (7) 知道常用的统计软件。

在当今的时代，当你浏览报纸、杂志或者网络时，会发现大量的信息。如何判断各种报道里根据数据所得出的结论是否有道理？样本是否具有代表性？样本是如何抽取的？很多人已经意识到，作为一个信息接收者或者使用者，在日常生活中，对统计学基本概念的了解以及懂得如何用统计去解决问题是非常重要的。正如书名所示，本书可以帮助你理解统计学，熟悉统计语言，并知道如何评估统计结果。如果你想研究统计学，本书将成为你开始漫长却又乐趣无穷的“统计”旅行的一个合适的起点。

为帮助你漫游统计乐园，本书的每一章都是以一些和本章内容有关的实际问题开始的。我们希望这些问题成为你对每套统计大餐的开胃菜。下面就以一些问题来开始第1章。

(1) 数据显示，2010年中国名义国内生产总值（GDP）为58 786亿美元，GDP增速为10%左右。而2010年日本名义GDP为54 742亿美元，GDP增速为3.9%。中国GDP比日本多4 044亿美元，这意味着中国超越日本成为世界第二大经济体。尽管中国总量GDP超过日本，但中国人均GDP却只有日本的1/10，在世界排名第100位左右，不到世界平均水平的一半（见图1.1）。这样的结果对我们意味着什么呢？毫无疑问，我们需要知道一些统计知识。

(2) 在金融危机发生后，各国政府对于国家经济的现状和对未来经济走势的预测都十分关注，每个月都要编制消费价格指数（CPI）等各类指数以评价通货膨胀情况。有关商品销售额、新开工的住宅、货币流通以及工业生产的信息仅仅是构成预测基础的成百上千类信息的一小部分。该如何解读这些数据呢？知道一些统计知识有助于我们知道这些数据是如何与我们的生活息息相关的。

(3) 作为一名潜在的汽车购买者和一名有责任心的市民，你乐意为保护地球生态环境做出应有的贡献。根据最新的研究成果，消费者的行为对自然资源到底有怎样的影响呢？你应该购买柴油发动机的汽车，还是购买电动车，或者干脆骑自行车呢？报纸、杂志或网络上的统计结果对于你的决定非常关键，这些结果会影响你该有怎样的消费观。

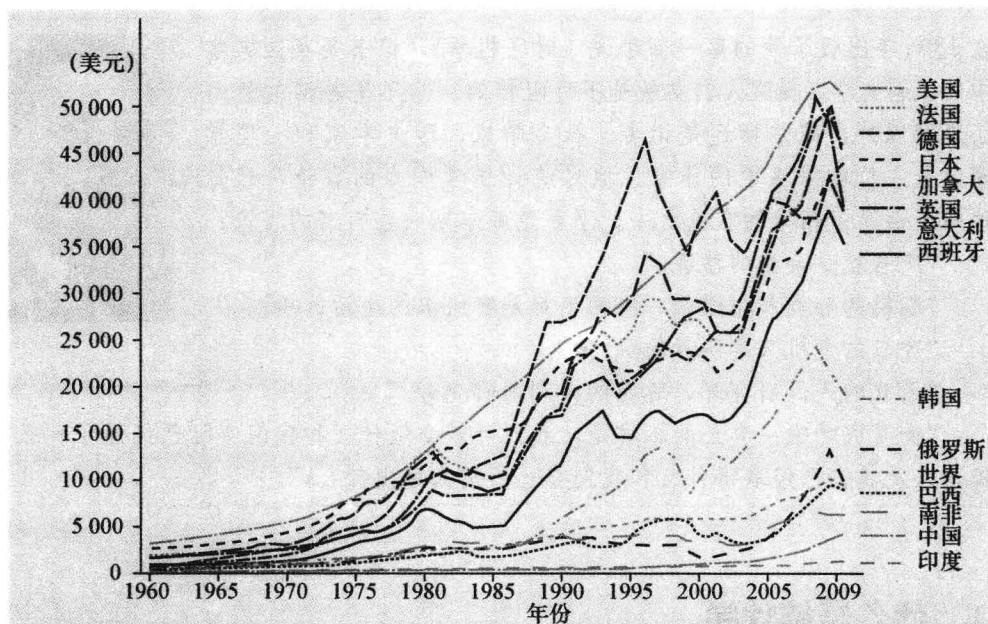


图 1.1 1960—2009 年全球主要国家人均 GDP 的增长图表

资料来源：Google 网站。World Bank, World Development Indication.



低碳社会要求汽车厂商根据它们的总产量生产一定比例的电动车，以减少机动车造成的空气污染，这已开始成为全世界的趋势。对于立法者来说，统计信息在使他们相信并检测使用电动车在改进空气质量上的有效性方面起着关键作用。

资料来源：[http://car.auto.ifeng.com/photo/carpiclarge/2171/1100323#pic\\_nav](http://car.auto.ifeng.com/photo/carpiclarge/2171/1100323#pic_nav).

(4) 假设你是市场部的新任经理，有一项广告活动的统计结果摆到了你面前，声称某个结果是“统计显著”的。你如何解释这份报告而又不暴露你对该术语的无知呢？赶快学点统计，这对你和你的事业都非常有用。

经典科幻小说《星际战争》和《时间机器》的作者威尔斯曾说过：“终有一天，统计思考的能力将会和读写的能力一样重要。”威尔斯 100 多年前的预言，在今天变成了现实。

### ☆人物小传

#### 赫伯特·乔治·威尔斯

出生于英国的作家赫伯特·乔治·威尔斯 (Herbert George Wells, 1866—1946)



在 1895 年出版了他的第一部小说《时间机器》。该书是英国文学中的讽刺文学，提醒人们发展是不可阻挡的。除了是著名的科幻小说家以外，威尔斯还是记者、社会学家、历史学家和心理学家，是一位硕果累累的作家。他对统计思考能力的言论，只是他对生活众多精辟的言论之一。下面是威尔斯一些有名的言论：

- “广告就是合法的撒谎。”
- “粗糙的分类和错误的一概而论都是有组织的生活的祸端。”
- “今日的危机只是明天的玩笑。”
- “愚蠢的人做研究者，聪明的人利用研究者。”
- “衡量成功唯一真正的标准是，我们可能会做什么和我们实际已经做了什么的比较，以及我们已经取得什么和我们自己为别人做出什么。”



## 1.1 什么是统计学

在 300 多年前，统计的英文单词 Statistics 首次被人们应用，指政府部门记录人们出生和死亡信息的工作。时至今日，统计仍然是世界上各个层次的政府机构的支柱。全球统计已成为许多国际组织，像跨国公司、联合国和一些关注人口密度、生态灾难和疾病流行的组织等重点关注的对象。2010 年 6 月 3 日，第 64 届联合国大会第 90 次会议通过决议，确定每年 10 月 20 日为“世界统计日”。2011 年初，国务院学位委员会在新的研究生专业目录中将统计学上升为一级学科，为统计学科和统计教育的发展提供了更广阔的平台。同时，也凸显出统计对科学研究、经济增长和社会发展的重要性。

Statistics 这个单词通常有两个含义：当它以单数名词出现时，表示一门学科的名称“统计学”；当它以复数名词出现时，表示“统计数据”或“统计资料”，此时，统计可以认为是某种形式的数字，例如北京的月平均降水量、上海的周平均气温、中国持有的外汇存款余额等。我们面对着所谓的“信息爆炸”，数字化的信息无处不在。现代社会似乎对统计有一种永不满足的需求。正是基于这种需求，统计学家收集越来越多的统计数字。统计数字又叫数据。Statistics 的一个简单意思就是数值数据。在本书中，我们超越了这种意思。我们感兴趣的是如何收集数据和处理数据。最后，我们希望通过统计的帮助把数据中的信息变成实际的知识。

简而言之，统计学是研究收集数据、整理数据、分析数据以及由数据分析得出结论的方法，简称为“数据的科学”。统计学有时候也指一门学科，例如你正在通过本教材学习的统计学课程。在统计学中，统计学家探索、开创各种收集数据和分析数据的方法以得出结论。他们从数学公式中为统计设计出新的应用，并通过实践来检验理论模型。虽然统计学没有自己的基于实验或观察的经验研究对象，它不像

化学、社会学或心理学这类学科研究已经定义得很好的某种现象，但是，统计学为经济学家、社会学家、化学家、心理学家以及其他人提供了一套研究对象的方法。

### ☆人物小传

#### 瓦尔德

有人把统计学定义为数据处理的一门艺术，瓦尔德(A. Wald, 1902—1950)的事例就是明证。瓦尔德是第二次世界大战时期的统计学家，他发明的一些统计方法在战时被视为军事机密。当有人向瓦尔德咨询飞机上什么部位应该加强钢板时，他开始研究从战役中返航的军机上受敌军创伤的弹孔位置。他画了飞机的轮廓，并且标示出弹孔的位置。资料累积一段时间后，几乎把机身各部位都填满了。于是瓦尔德提议，把剩下少数几个没有弹孔的部位补强……因为这些部位被击中的飞机都没有返航。这是一个简单但近乎完美的实例，简单的统计方法一旦融入了统计学家的智慧，便显得生动而唯美！



我们希望你学习本书之后，能够通过统计学的帮助欣赏到数据是如何转化为比数字本身更为复杂、更为有用的知识的。

## 1.2 如何运用统计学

统计在许多学科中得到广泛的应用，可以毫不夸张地说，我们已经被统计分析结果包围。比如说，学术研究杂志就依赖于统计结果。在许多学科中，一篇文章是否能够发表在主要杂志上，在很大程度上取决于该文章是否能正确地使用统计方法。

除了学术领域，统计在其他领域也广泛使用。统计在工业中得到大量使用，尤其用于研究新产品、质量控制和市场开发中；市场调查也需要统计。在当今时代，很难想象不需要市场调查来获知销售者对于企业产品或服务的看法。产品或服务的改进依赖于由统计得到的消费者的反馈信息；同样，没有一份报纸、杂志不刊登以统计为基础的文章。我们在报纸、杂志、网络等媒介上可以读到诸如在2010年上海世博会期间参观的人数，为汶川和舟曲等灾区捐款的金额，大学生就业率等统计数据；甚至在电视节目中，主持人的出现和电视剧的播放都依赖于统计，因为只有排名靠前受观众喜爱的主持人和节目才能在竞争中生存。



### 想一想 1.1

请从你的实际经验中想出一些以统计分析为决策依据的例子，并思考出现这种统计结果的原因。

学习统计可以帮助我们理解什么可能出差错。作为信息接收者或者使用者，如果我们希望能完全理解统计的广泛应用，就应知道为得到所读到的结果而需要的原则和方法。对统计的了解可以帮助我们评价这些结果，并使我们随时保持警觉，因为从开始叙述问题到完成最终报告的过程中随时可能出现一些错误。

## 《读者文摘》的总统选举调查

这个故事有关一次著名的失败的统计调查，它一直是一个统计逸事。在 1936 年美国总统选举前，一份名为《读者文摘》(Literary Digest) 的颇受人尊重的杂志进行了一次民意调查。调查的焦点当然是谁将成为下一届总统——是挑战者，堪萨斯州州长阿尔夫·兰登 (Alf Landon)，还是现任总统富兰克林·德拉诺·罗斯福 (Franklin Delano Roosevelt)。为了解选民意向，民意调查专家根据电话簿和车辆登记簿上的名单给一大批人发了简单的调查表（电话和汽车在 1936 年并不像现在这样普遍，但是这些名单比较容易得到）。尽管发出的调查表大约有 1 000 万份，但收回的比例并不高。在收回的调查表中，兰登非常受欢迎。于是，该杂志预测兰登将赢得选举。

如果你有一些统计知识，你会对这个声称兰登将赢得选举的预测结果怀有疑问。正如你所怀疑的，在经济大萧条时期调查拥有电话和汽车的人，并不能够很好地反映全体选民的观点。此外，只有少数的调查表被收回，这一点也是值得怀疑的。事实表明，最终是罗斯福而不是兰登赢得了这次选举。由此可见，该杂志的调查结果有多么错误。当前大多数应用统计不会像这个例子错得那样离谱，但即便在今天，我们也很容易发现统计被误用的情况，尤其在需要考虑选择正确的样本时。

需要提及的一点是，当时盖洛普 (Gallup) 也进行了调查，它的调查结果是罗斯福获胜。在这次总统选举之后，盖洛普成了一家比较成功的调查公司。

我们还需要理解一些统计术语。如果不理解那些统计术语的话，统计分析结果就不可能给我们提供太多帮助。例如，统计报告中最典型的一个术语是“统计显著”；在给出选民偏爱某一候选人的比例时，可能会有“样本误差等于±3%”或“边际误差等于±3%”等术语；两个变量可能“高度相关”。以上是三种最常见的统计术语。对于知道它们的人来说，这些术语包含了有用的信息。而不知道它们意义的人则有可能根本不知道它们在说什么，甚至得出错误的结论。