



“十二五”普通高等教育本科国家级规划教材
北京高等教育精品教材



教育部经济管理类核心课程教材

Statistics

统计学

(第7版)

▶▶ 贾俊平 编著

“十二五”普通高等教育本科国家级规划教材
北京高等教育精品教材



教育部经济管理类核心课程教材



Statistics

统计学

(第7版)

▶▶ 贾俊平 编著

中国人民大学出版社
· 北京 ·

图书在版编目 (CIP) 数据

统计学/贾俊平编著. —7 版. —北京: 中国人民大学出版社, 2018. 5
 教育部经济管理类核心课程教材
 ISBN 978-7-300-25687-0

I. ①统… II. ①贾… 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 (2018) 第 069304 号

“十二五”普通高等教育本科国家级规划教材 北京高等教育精品教材 教育部经济管理类核心课程教材 统计学 (第 7 版) 贾俊平 编著 Tongjixue		出版发行 中国人民大学出版社 社 址 北京中关村大街 31 号 电 话 010-62511242 (总编室) 010-82501766 (邮购部) 010-62515195 (发行公司) 网 址 http://www.crup.com.cn http://www.ttrnet.com (人大教研网) 经 销 新华书店 印 刷 北京密兴印刷有限公司 规 格 185 mm×260 mm 16 开本 印 张 18.75 插页 1 字 数 400 000	邮 政 编 码 100080 010-62511770 (质管部) 010-62514148 (门市部) 010-62515275 (盗版举报)	版 次 2003 年 6 月第 1 版 2018 年 5 月第 7 版 印 次 2018 年 6 月第 2 次印刷 定 价 39.00 元
--	--	---	---	--

版权所有 侵权必究

印装差错 负责调换

内 容 简 介

统计学作为研究数据的一门科学，为使用者提供了一套获取数据、分析数据并从数据中得出结论的原则和方法。掌握统计技术可以让你在竞争激烈的就业市场中占据优势。本书在第6版的基础上修订而成，主要有以下特色：

(1) 强调对统计思想的阐述，通过实例讲解统计方法，强调与计算机的结合。

(2) 对例题和习题数据进行了部分更新，所有实例均给出SPSS和Excel的实现过程和结果。

(3) 考虑到教学时数限制，删除了一些章节，在保持体系完整性的同时，内容更加简洁易懂。

(4) 配有内容丰富的教学资源库，便于读者学习。



作者简介

贾俊平 中国人民大学统计学院副教授。研究方向：统计方法在经济各领域的应用，统计教学方式和方法。主要著作有：《统计学——基于R》《统计学——基于SPSS》《统计学——基于Excel》。主持研究的“非统计学专业本科公共基础课——统计学教学改革”项目获2001年国家级教学成果二等奖、2001年北京市教学成果一等奖。2001年荣获北京市经济技术创新标兵称号，2003年荣获宝钢优秀教师奖等。

策划编辑 王伟娟
责任编辑 胡志敏
封面设计 三众工作室 / 耿中虎
版式设计 赵星华

教育部经济管理类核心课程教材

出版说明

前言

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 and 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

统计作为数据分析的一种通用语言，已在越来越多的领域加以应用。对很多人而言，掌握统计技术可在竞争日趋激烈的就业市场中占据优势。统计学作为研究数据的一门科学，为使用者提供了一套获取数据、分析数据并从数据中得出结论的原则和方法。

多数人都把统计作为一门难学的课程来看待，其实统计并不像人们想象的那么难，关键是看你怎么学。如果在学习过程中把注意力放在公式上，放在数据的计算过程上，而忽视对统计思想的理解，不仅难以学会，也难以将统计用到实处。如果抛开复杂公式的表象，把繁杂的计算过程交给计算机来完成，就会发现统计其实很简单。统计的精髓是使复杂问题简单化，而不是把简单问题复杂化；统计的真谛在于它所体现的思想，在于它所提供的思维方式；学好统计的关键是掌握如何运用统计思维来思考问题，而不是简单地记住那些死的统计知识。

本书具有以下特点：

第一，首次将非参数检验纳入《统计学》教材。

第二，实现教材内容与计算机的完全结合。除部分为展示方法的计算过程外，多数统计计算都实现了计算机化。在讲清楚统计方法原理和思想的基础上，基本上都给出了由统计软件实现计算的详细操作步骤，并对输出结果做了详细解读。考虑到读者对统计软件的可获得性和熟悉程度，本书结合使用了 SPSS (19.0 中文版) 和 Excel (2013 版) 两个软件。其中大部分方法使用 SPSS 实现计算和分析，少量使用 Excel。对每种软件都给出了具体的操作步骤，读者按此操作就会得到所需的统计分析结果。

第三，本书配有内容丰富的教学和学习资源库，内容包括教学和学习用 PPT、教材例题和练习题的数据文件。此外，本书配套的学习指导书给出了每一章的知识结构、学习要点、选择题和答案，以及教材后所附习题的详细解答。

本书可作为高等院校经济管理类专业本科生统计学课程的教材，也可作为研究生和 MBA 的教材或参考书。在使用中，教师对有些章节可根据教学需要和教学时数



酌情选讲。希望本书能对您有所帮助,也希望您提出更多的修改建议,以便进一步修改和完善。

贾俊平

于中国人民大学统计学院

前 言

本书是根据教育部颁布的《统计学专业本科课程方案》和《统计学专业本科课程教学大纲》的要求,在广泛吸收国内外优秀教材的基础上,结合我国统计学教学改革的实际编写而成的。本书可作为高等院校统计学专业及相关专业的教材,也可供从事统计工作的工程技术人员参考。

本书在编写过程中,参考了国内外许多优秀的统计学教材和有关文献,并得到了许多统计学专家的指导和帮助,在此表示衷心的感谢。本书在编写过程中,得到了中国人民大学统计学院领导和同事们的支持,在此表示衷心的感谢。本书在编写过程中,得到了许多统计学专家的指导和帮助,在此表示衷心的感谢。

本书在编写过程中,参考了国内外许多优秀的统计学教材和有关文献,并得到了许多统计学专家的指导和帮助,在此表示衷心的感谢。本书在编写过程中,得到了中国人民大学统计学院领导和同事们的支持,在此表示衷心的感谢。本书在编写过程中,得到了许多统计学专家的指导和帮助,在此表示衷心的感谢。

本书在编写过程中,参考了国内外许多优秀的统计学教材和有关文献,并得到了许多统计学专家的指导和帮助,在此表示衷心的感谢。本书在编写过程中,得到了中国人民大学统计学院领导和同事们的支持,在此表示衷心的感谢。本书在编写过程中,得到了许多统计学专家的指导和帮助,在此表示衷心的感谢。

第7版的变化

本书在第6版的基础上修订而成。与第6版相比，第7版主要有以下变化：

第一，Excel软件的使用版本由Excel 2010升级为Excel 2013。书中所用的Excel函数及其他功能也均按Excel 2013做了修改。

第二，增加了附录2：SPSS和Excel中的常用统计函数，列出了函数表达式、参数的含义及返回结果等，供读者查阅和使用。

第三，第3章增加了百分位数的计算和应用，并对部分例题数据做了更新。

第四，考虑到教学时数的限制，删除了第6版中的第12章（主成分分析和因子分析）、第13章（聚类分析），删除了第6版中第11章（时间序列预测）11.3节中的11.3.3（残差自相关及其检验）、11.5节（Box-Jenkins方法：ARIMA模型）。为保持教材的完整性，仍然保留了非参数检验一章的内容，以便读者了解参数检验不适用情形下的统计处理方法。

本书可作为一个学期3~4学分统计学课程的教材使用，教师对有些内容可根据教学时数、教学对象的特点等酌情选讲。

软件说明	11
参考文献	27
第2章 用图表展示数据	13
2.1 饼图系列，用圆饼图形表示数据	14
2.2 用图表展示定性数据	15
2.2.1 生成频数分布表	15
2.2.2 定性数据的展示	16
2.3 用图表展示定量数据	21
2.3.1 生成频数分布表	21
2.3.2 定量数据的展示	23
2.4 在图例中使用表	22
本章小结——数据展示与图例展示方法	33
习题	32

目 录

第 1 章 统计、数据和计算机	1
开篇案例：怎样理解这样一些统计结论？	1
1.1 统计及其应用领域	2
1.1.1 统计学研究什么	2
1.1.2 统计的应用	3
1.2 怎样获得统计数据	5
1.2.1 变量与数据	5
1.2.2 数据的来源	7
1.3 统计与计算机	8
本书图解：统计方法分类与本书框架	10
主要术语	11
软件应用	11
思考与练习	12
第 2 章 用图表展示数据	14
开篇案例：用哪些图形展示奖牌？	14
2.1 用图表展示定性数据	15
2.1.1 生成频数分布表	15
2.1.2 定性数据的图示	18
2.2 用图表展示定量数据	21
2.2.1 生成频数分布表	21
2.2.2 定量数据的图示	23
2.3 合理使用图表	32
本章图解：数据类型与图表展示方法	33
主要术语	33



软件应用	34
思考与练习	36
第3章 用统计量描述数据	38
开篇案例: 哪名运动员的发挥更稳定?	38
3.1 水平的度量	39
3.1.1 平均数	39
3.1.2 分位数	40
3.1.3 用哪个值代表一组数据	43
3.2 差异的度量	44
3.2.1 极差和四分位差	44
3.2.2 方差和标准差	44
3.2.3 比较几组数据的离散程度: 离散系数	47
3.3 分布形状的度量	48
本章图解: 数据分布特征与描述统计量	49
主要术语	50
软件应用	50
思考与练习	51
第4章 概率分布	53
开篇案例: 神舟七号飞船遭遇空间碎片的概率有多大?	53
4.1 度量事件发生的可能性	54
4.2 随机变量的概率分布	55
4.2.1 随机变量及其概括性度量	55
4.2.2 离散型概率分布	57
4.2.3 连续型概率分布	59
4.3 其他几个重要的统计分布	61
4.3.1 t 分布	61
4.3.2 χ^2 分布	62
4.3.3 F 分布	63
4.4 样本统计量的概率分布	63
4.4.1 统计量及其分布	63
4.4.2 样本均值的分布	64
4.4.3 其他统计量的分布	67
4.4.4 统计量的标准误差	67
本章图解: 随机变量的概率分布	68
主要术语	68
软件应用	69

思考与练习	70
第 5 章 参数估计	72
开篇案例：大学生每周上网花多少时间？	72
5.1 参数估计的基本原理	73
5.1.1 点估计与区间估计	73
5.1.2 评价估计量的标准	76
5.2 一个总体参数的区间估计	78
5.2.1 总体均值的区间估计	78
5.2.2 总体比例的区间估计	80
5.2.3 总体方差的区间估计	81
5.3 两个总体参数的区间估计	82
5.3.1 两个总体均值之差的区间估计	82
5.3.2 两个总体比例之差的区间估计	87
5.3.3 两个总体方差比的区间估计	87
5.4 样本量的确定	88
5.4.1 估计总体均值时样本量的确定	89
5.4.2 估计总体比例时样本量的确定	90
本章图解：参数估计所使用的分布	91
主要术语	92
软件应用	92
思考与练习	93
第 6 章 假设检验	96
开篇案例：正常人的平均体温是 37°C 吗？	96
6.1 假设检验的基本原理	97
6.1.1 怎样提出假设	97
6.1.2 怎样作出决策	99
6.1.3 怎样表述决策结果	103
6.2 一个总体参数的检验	104
6.2.1 总体均值的检验	104
6.2.2 总体比例的检验	108
6.2.3 总体方差的检验	108
6.3 两个总体参数的检验	110
6.3.1 两个总体均值之差的检验	110
6.3.2 两个总体比例之差的检验	113
6.3.3 两个总体方差比的检验	115
6.4 总体分布的检验	117

6.4.1	正态性检验的图示法	117
6.4.2	Shapiro-Wilk 和 K-S 正态性检验	118
	本章图解: 假设检验的内容框架	121
	主要术语	122
	软件应用	122
	思考与练习	124
第7章	类别变量的推断	127
	开篇案例: 性别与是否逃课有关系吗?	127
7.1	一个类别变量的拟合优度检验	128
7.1.1	期望频数相等	128
7.1.2	期望频数不等	130
7.2	两个类别变量的独立性检验	131
7.2.1	列联表与 χ^2 独立性检验	132
7.2.2	应用 χ^2 检验应注意的问题	134
7.3	两个类别变量的相关性度量	134
7.3.1	ϕ 系数和 Cramer's V 系数	134
7.3.2	列联系数	135
	本章图解: 类别变量检验方法	136
	主要术语	136
	软件应用	136
	思考与练习	137
第8章	方差分析与实验设计	140
	开篇案例: 不同运动队的平均成绩之间是否有显著差异?	140
8.1	方差分析的基本原理	141
8.1.1	什么是方差分析	141
8.1.2	误差分解	142
8.2	单因素方差分析	143
8.2.1	数学模型	144
8.2.2	效应检验	145
8.2.3	多重比较	148
8.3	双因素方差分析	152
8.3.1	数学模型	152
8.3.2	只考虑主效应	154
8.3.3	考虑交互效应	159
8.4	方差分析的假定及其检验	162
8.4.1	正态性检验	162

8.4.2	方差齐性检验	163
8.5	实验设计初步	166
8.5.1	完全随机化设计	166
8.5.2	随机化区组设计	167
8.5.3	析因设计	168
	本章图解：方差分析过程	170
	主要术语	170
	软件应用	171
	思考与练习	173
第9章	一元线性回归	176
	开篇案例：子代与父代一样吗？	176
9.1	变量间的关系	177
9.1.1	变量间是什么样的关系	177
9.1.2	用散点图描述相关关系	178
9.1.3	用相关系数度量关系强度	179
9.2	一元线性回归模型的估计和检验	182
9.2.1	一元线性回归模型	182
9.2.2	参数的最小二乘估计	183
9.2.3	模型的拟合优度	186
9.2.4	显著性检验	188
9.3	利用回归方程进行预测	190
9.3.1	平均值的置信区间	190
9.3.2	个别值的预测区间	191
9.4	用残差检验模型的假定	193
9.4.1	检验方差齐性	193
9.4.2	检验正态性	195
	本章图解：一元线性回归的建模过程	197
	主要术语	197
	软件应用	198
	思考与练习	199
第10章	多元线性回归	202
	开篇案例：身高受哪些因素影响？	202
10.1	多元线性回归模型	203
10.1.1	回归模型与回归方程	203
10.1.2	参数的最小二乘估计	205
10.2	拟合优度和显著性检验	207

10.2.1	模型的拟合优度	207
10.2.2	显著性检验	209
10.3	多重共线性及其处理	210
10.3.1	多重共线性及其识别	210
10.3.2	变量选择与逐步回归	212
10.4	利用回归方程进行预测	215
10.5	哑变量回归	216
10.5.1	在模型中引进哑变量	217
10.5.2	含有一个哑变量的回归	217
	本章图解:多元线性回归的建模过程	222
	主要术语	222
	软件应用	223
	思考与练习	224
第 11 章	时间序列预测	227
	开篇案例:下个月的消费者信心指数是多少?	227
11.1	时间序列的成分和预测方法	228
11.1.1	时间序列的成分	228
11.1.2	预测方法的选择与评估	231
11.2	平稳序列的预测	232
11.2.1	移动平均预测	232
11.2.2	简单指数平滑预测	232
11.3	趋势预测	234
11.3.1	线性趋势预测	234
11.3.2	非线性趋势预测	238
11.4	多成分序列的预测	241
11.4.1	Winter 指数平滑预测	241
11.4.2	引入季节哑变量的多元回归预测	242
11.4.3	分解预测	246
	本章图解:时间序列预测的程序和方法	248
	主要术语	249
	软件应用	249
	思考与练习	250
第 12 章	非参数检验	254
	开篇案例:不同商圈的报纸发行量是否有差异?	254
12.1	单样本的检验	255
12.1.1	中位数的符号检验	255



12.1.2 Wilcoxon 符号秩检验	257
12.2 两个及两个以上样本的检验	259
12.2.1 两个配对样本的 Wilcoxon 符号秩检验	259
12.2.2 两个独立样本的 Mann-Whitney 检验	261
12.2.3 k 个独立样本的 Kruskal-Wallis 检验	263
12.3 秩相关及其检验	265
12.3.1 Spearman 秩相关及其检验	266
12.3.2 Kendall 秩相关及其检验	267
本章图解：非参数检验方法	270
主要术语	270
软件应用	270
思考与练习	271
附录 1 解读指数	274
附录 2 SPSS 和 Excel 中的常用统计函数	277
参考文献	281

统计、数据和计算机

统计思维总有一天会像读与写一样成为一个有效率公民的必备能力。

——H. G. Wells

开篇案例：怎样理解这样一些统计结论？

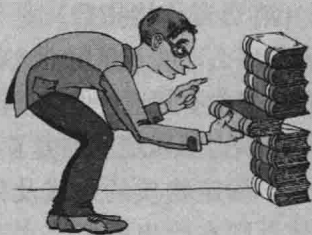
每天你都会看到各种统计数字，但你或许没有仔细想过它们意味着什么。看看下面的一些统计研究结果，你会有怎样的看法呢？

- 吸烟对健康是有害的，吸香烟的男性减少寿命 2 250 天。
- 不结婚的男性会减少寿命 3 500 天，不结婚的女性会减少寿命 1 600 天。
- 身体超重 30% 会使寿命减少 1 300 天。
- 每天摄取 500 毫升维生素 C，生命可延长 6 年。
- 身材高的父亲，其子女的身材也较高。
- 第二个出生的子女没有第一个聪明，第三个出生的子女没有第二个聪明，依此类推。

● 学生们在听了莫扎特钢琴曲 10 分钟后的推理，要比他们听 10 分钟娱乐性的其他曲目后的推理做得更好。

● 漂亮的女性有损男性的智力。男性在看到漂亮女性时智力会下降，这就是为什么大学里的女孩子比男孩子学习好的原因。

● 上课坐在前面的学生平均考试分数比坐在后面的学生高。



● 中国科学院空间环境研究预报中心的专家称，在神舟七号载人航天飞行期间，飞船遭遇空间碎片的概率在百万分之一以下。

看懂这些结论并不困难，但这些结论是怎样得出来的？你相信这些结论吗？学点儿统计学知识你就会正确理解它们。