



“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

(第六版)

▶▶ 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

(第六版)

▶▶ 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. —6 版. —北京: 中国人民大学出版社, 2016. 6
教育部经济管理类核心课程教材
ISBN 978-7-300-23041-2

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 (2016) 第 137116 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 (第六版)
贾俊平 编著
Tongjixue

出版发行	中国人民大学出版社	邮政编码	100080
社 址	北京中关村大街 31 号		
电 话	010-62511242 (总编室)	010-62511770 (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	版 次	2003 年 6 月第 1 版
印 刷	北京密兴印刷有限公司		2016 年 6 月第 6 版
规 格	185 mm×260 mm 16 开本	印 次	2016 年 8 月第 2 次印刷
印 张	21.5 插页 1	定 价	39.00 元
字 数	474 000		

版权所有 侵权必究 印装差错 负责调换

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

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，首次将部分多元统计方法和非参数检验纳入《统计学》教材。本书包括一元线性回归、多元线性回归、主成分分析和因子分析、聚类分析以及非参数检验等。计算机应用的普及，使这些方法的实际应用成为可能。

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 (2010 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

第六版的变化

与第五版相比，第六版的变化主要有以下几个方面：

第1章“统计、数据和计算机”：1.3节中删除了S-PLUS的介绍，增加了R软件的介绍，因为R已逐步成为数据分析软件的主流。在软件应用中给出了Excel 2010版中【数据分析】工具安装的步骤。

第2章“用图表展示数据”：在软件应用部分给出了Excel 2010版做数据透视表的操作步骤，替换了2003版的操作步骤。

第4章“概率分布”：删除了第五版4.2.3节中数据的正态性评估内容；在软件应用部分删除了用SPSS绘制正态概率图的操作步骤。

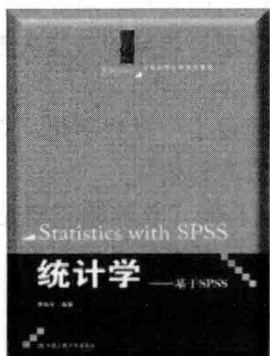
第6章“假设检验”：增加了一节内容，即“6.4总体分布的检验”；介绍了正态性检验的图示法，如P-P图和Q-Q图，以及Shapiro-Wilk和K-S正态性检验。

第8章“方差分析与实验设计”：删除了第五版中的“8.1.3方差分析的基本假定”；重新编写了“8.2.3多重比较”，给出Fisher多重比较的LSD方法的检验统计量和差值置信区间的计算方法，同时，增加了Tukey-Kramer的HSD方法；增加了一节内容，即“8.4方差分析的假定及其检验”，介绍了方差分析中正态性和方差齐性的检验方法。

第10章“多元线性回归”：增加了标准化回归系数的解释。

第14章“非参数检验”：删除了第五版中的“14.1.1总体分布类型的检验”，将相应内容放到了第6章假设检验中的6.4节。

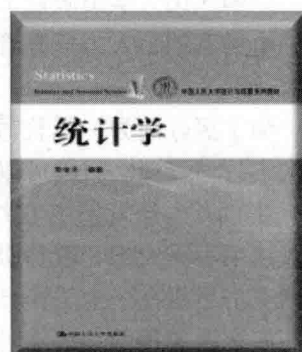
贾俊平的其他相关著作



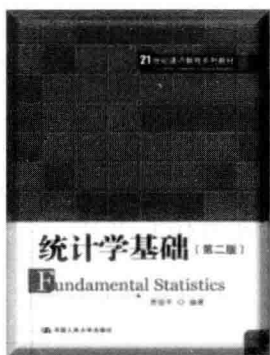
本书基于 SPSS 实现全部例题的计算与分析，并给出了 SPSS 的详细操作步骤和结果。内容包括数据的描述性分析方法、推断方法以及其他常用的一些统计方法等。可作为高等院校经济管理类专业以及部分理、工、农、林、医、药专业的本科生教材使用。



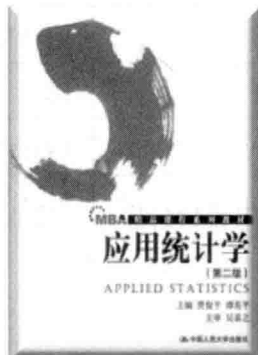
本书基于 R 实现全部例题的计算与分析，并给出了 R 的详细程序和结果。内容包括数据的描述性分析方法、推断方法以及其他常用的一些统计方法等。可作为高等院校经济管理类专业以及部分理、工、农、林、医、药专业的本科生教材使用。



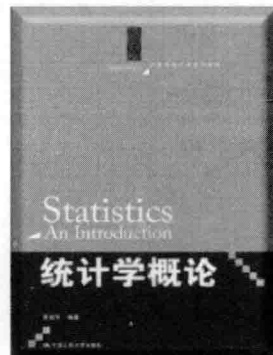
本书基于 SPSS 和 R 实现例题计算与分析，并给出了 SPSS 操作步骤和 R 程序。内容包括描述统计、推断统计、非参数检验以及实际中常用的一些统计方法等。可作为高等院校统计学专业本科生教材使用，也可作为其他专业本科生和研究生的教材使用。



本书包括描述统计、推断统计及一些常用的基本统计方法。使用 Excel 软件实现例题计算与分析。可作为高等院校的全校通识课教材，也可作为一般性院校的本科生教材使用。



本书是为 MBA 专门编写的统计教材。内容包括描述统计、推断统计，多元统计和非参数检验等统计方法。结合 SPSS 和 Excel 实现例题的计算与分析。可作为 MBA 和 EMBA 等的教材使用。



本书是为统计学专业的本科生而编写的学科基础课教材。从实际问题入手介绍统计方法的思想和应用，为其他课程的学习奠定基础。结合使用 SPSS 和 Excel 两个软件。

教师教学服务说明

中国人民大学出版社工商管理分社以出版经典、高品质的工商管理、财务会计、统计、市场营销、人力资源管理、运营管理、物流管理、旅游管理等领域的各层次教材为宗旨。

为了更好地为一线教师服务，近年来工商管理分社着力建设了一批数字化、立体化的网络教学资源。教师可以通过以下方式获得免费下载教学资源的权限：

在“人大经管图书在线”（www.rdjg.com.cn）注册，下载“教师服务登记表”，或直接填写下面的“教师服务登记表”，加盖院系公章，然后邮寄或传真给我们。我们收到表格后将在一个工作日内为您开通相关资源的下载权限。

如您需要帮助，请随时与我们联系：

中国人民大学出版社工商管理分社

联系电话：010-62515735, 62515749, 62515987

传 真：010-62515732, 62514775 电子邮箱：rdcbsjg@crup.com.cn

通讯地址：北京市海淀区中关村大街甲 59 号文化大厦 1501 室（100872）

教师服务登记表

姓名		<input type="checkbox"/> 先生 <input type="checkbox"/> 女士	职 称		
座机/手机			电子邮箱		
通讯地址			邮 编		
任教学校			所在院系		
所授课程	课程名称	现用教材名称	出版社	对象（本科生/研究生/MBA/其他）	学生人数
需要哪本教材的配套资源					
人大经管图书在线用户名					
院/系领导（签字）： 院/系办公室盖章					



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

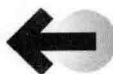
软件应用	34
思考与练习	36
第3章 用统计量描述数据	38
开篇案例：哪名运动员的发挥更稳定？	38
3.1 水平的度量	39
3.1.1 平均数	39
3.1.2 中位数和分位数	40
3.1.3 用哪个值代表一组数据	42
3.2 差异的度量	43
3.2.1 极差和四分位差	43
3.2.2 方差和标准差	43
3.2.3 比较几组数据的离散程度：离散系数	46
3.3 分布形状的度量	47
本章图解：数据分布特征与描述统计量	48
主要术语	49
软件应用	49
思考与练习	50
第4章 概率分布	52
开篇案例：神舟七号飞船遭遇空间碎片的概率有多大？	52
4.1 度量事件发生的可能性	53
4.2 随机变量的概率分布	54
4.2.1 随机变量及其概括性度量	54
4.2.2 离散型概率分布	56
4.2.3 连续型概率分布	58
4.3 其他几个重要的统计分布	59
4.3.1 t 分布	60
4.3.2 χ^2 分布	60
4.3.3 F 分布	61
4.4 样本统计量的概率分布	62
4.4.1 统计量及其分布	62
4.4.2 样本均值的分布	63
4.4.3 其他统计量的分布	65
4.4.4 统计量的标准误差	66
本章图解：随机变量的概率分布	67
主要术语	67
软件应用	68

思考与练习	69
第 5 章 参数估计	71
开篇案例：大学生每周上网花多少时间？	71
5.1 参数估计的基本原理	72
5.1.1 点估计与区间估计	72
5.1.2 评价估计量的标准	75
5.2 一个总体参数的区间估计	77
5.2.1 总体均值的区间估计	77
5.2.2 总体比例的区间估计	79
5.2.3 总体方差的区间估计	80
5.3 两个总体参数的区间估计	81
5.3.1 两个总体均值之差的区间估计	81
5.3.2 两个总体比例之差的区间估计	86
5.3.3 两个总体方差比的区间估计	86
5.4 样本量的确定	87
5.4.1 估计总体均值时样本量的确定	88
5.4.2 估计总体比例时样本量的确定	89
本章图解：参数估计所使用的分布	90
主要术语	91
软件应用	91
思考与练习	92
第 6 章 假设检验	95
开篇案例：正常人的平均体温是 37°C 吗？	95
6.1 假设检验的基本原理	96
6.1.1 怎样提出假设	96
6.1.2 怎样作出决策	98
6.1.3 怎样表述决策结果	102
6.2 一个总体参数的检验	103
6.2.1 总体均值的检验	103
6.2.2 总体比例的检验	107
6.2.3 总体方差的检验	107
6.3 两个总体参数的检验	109
6.3.1 两个总体均值之差的检验	109
6.3.2 两个总体比例之差的检验	112
6.3.3 两个总体方差比的检验	114
6.4 总体分布的检验	116



6.4.1	正态性检验的图示法	116
6.4.2	Shapiro-Wilk 和 K-S 正态性检验	117
	本章图解: 假设检验的内容框架	120
	主要术语	121
	软件应用	121
	思考与练习	123
第 7 章	类别变量的推断	126
	开篇案例: 性别与是否逃课有关系吗?	126
7.1	一个类别变量的拟合优度检验	127
7.1.1	期望频数相等	127
7.1.2	期望频数不等	129
7.2	两个类别变量的独立性检验	130
7.2.1	列联表与 χ^2 独立性检验	131
7.2.2	应用 χ^2 检验应注意的问题	133
7.3	两个类别变量的相关性度量	133
7.3.1	ϕ 系数和 Cramer's V 系数	133
7.3.2	列联系数	134
	本章图解: 类别变量检验方法	135
	主要术语	135
	软件应用	135
	思考与练习	136
第 8 章	方差分析与实验设计	139
	开篇案例: 不同运动队的平均成绩之间是否有显著差异?	139
8.1	方差分析的基本原理	140
8.1.1	什么是方差分析	140
8.1.2	误差分解	141
8.2	单因素方差分析	142
8.2.1	数学模型	143
8.2.2	效应检验	144
8.2.3	多重比较	147
8.3	双因素方差分析	151
8.3.1	数学模型	151
8.3.2	只考虑主效应	153
8.3.3	考虑交互效应	158
8.4	方差分析的假定及其检验	161
8.4.1	正态性检验	161

8.4.2 方差齐性检验	162
8.5 实验设计初步	165
8.5.1 完全随机化设计	165
8.5.2 随机化区组设计	166
8.5.3 析因设计	167
本章图解：方差分析过程	169
主要术语	169
软件应用	170
思考与练习	172
第9章 一元线性回归	175
开篇案例：子代与父代一样吗?	175
9.1 变量间的关系	176
9.1.1 变量间是什么样的关系	176
9.1.2 用散点图描述相关关系	177
9.1.3 用相关系数度量关系强度	178
9.2 一元线性回归模型的估计和检验	181
9.2.1 一元线性回归模型	181
9.2.2 参数的最小二乘估计	182
9.2.3 回归直线的拟合优度	185
9.2.4 显著性检验	187
9.3 利用回归方程进行预测	189
9.3.1 平均值的置信区间	189
9.3.2 个别值的预测区间	190
9.4 用残差检验模型的假定	192
9.4.1 检验方差齐性	192
9.4.2 检验正态性	194
本章图解：一元线性回归的建模过程	196
主要术语	196
软件应用	197
思考与练习	198
第10章 多元线性回归	201
开篇案例：身高受哪些因素影响?	201
10.1 多元线性回归模型	202
10.1.1 回归模型与回归方程	202
10.1.2 参数的最小二乘估计	204
10.2 拟合优度和显著性检验	206



10.2.1	回归方程的拟合优度	206
10.2.2	显著性检验	208
10.3	多重共线性及其处理	209
10.3.1	多重共线性及其识别	209
10.3.2	变量选择与逐步回归	211
10.4	利用回归方程进行预测	214
10.5	哑变量回归	215
10.5.1	在模型中引进哑变量	216
10.5.2	含有一个哑变量的回归	216
	本章图解: 多元线性回归的建模过程	221
	主要术语	221
	软件应用	222
	思考与练习	223
第 11 章	时间序列预测	226
	开篇案例: 下个月的消费者信心指数是多少?	226
11.1	时间序列的成分和预测方法	227
11.1.1	时间序列的成分	227
11.1.2	预测方法的选择与评估	230
11.2	平稳序列的预测	231
11.2.1	移动平均预测	231
11.2.2	简单指数平滑预测	231
11.3	趋势预测	233
11.3.1	线性趋势预测	233
11.3.2	非线性趋势预测	237
11.3.3	残差自相关及其检验	240
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	243
11.4.3	分解预测	246
11.5	Box-Jenkins 方法: ARIMA 模型	248
11.5.1	自相关与自相关图	248
11.5.2	Box-Jenkins 方法的基本思想	250
11.5.3	ARIMA 模型的识别	251
	本章图解: 时间序列预测的程序和方法	262
	主要术语	263
	软件应用	263
	思考与练习	265

第 12 章 主成分分析和因子分析	268
开篇案例：因子分析得到的是什么？	268
12.1 主成分分析	269
12.1.1 主成分分析的基本原理	269
12.1.2 主成分分析的数学模型	270
12.1.3 主成分分析的步骤	271
12.2 因子分析	274
12.2.1 因子分析的意义和数学模型	274
12.2.2 因子分析的步骤	276
12.2.3 因子分析的应用	278
本章图解：主成分分析和因子分析的步骤	283
主要术语	283
软件应用	284
思考与练习	285
第 13 章 聚类分析	288
开篇案例：怎样把消费者分类？	288
13.1 聚类分析基本原理	289
13.1.1 什么是聚类分析	289
13.1.2 相似性的度量	290
13.2 层次聚类	291
13.2.1 层次聚类的两种方式	291
13.2.2 类间距离的计算方法	292
13.2.3 层次聚类的应用	292
13.3 K-均值聚类	297
13.3.1 K-均值聚类的基本过程	297
13.3.2 K-均值聚类的应用	297
13.3.3 使用聚类方法的注意事项	301
本章图解：聚类分析方法	302
主要术语	302
软件应用	302
思考与练习	303
第 14 章 非参数检验	305
开篇案例：不同商圈的报纸发行量是否有差异？	305
14.1 单样本的检验	306
14.1.1 中位数的符号检验	306
14.1.2 Wilcoxon 符号秩检验	308



14.2 两个及两个以上样本的检验	310
14.2.1 两个配对样本的 Wilcoxon 符号秩检验	310
14.2.2 两个独立样本的 Mann-Whitney 检验	312
14.2.3 k 个独立样本的 Kruskal-Wallis 检验	314
14.3 秩相关及其检验	316
14.3.1 Spearman 秩相关及其检验	317
14.3.2 Kendall 秩相关及其检验	318
本章图解: 非参数检验方法	321
主要术语	321
软件应用	321
思考与练习	322
附录 解读指数	325
参考文献	328