

● 高等学校教材

Statistics

统计学

管于华 主编



高等教育出版社
HIGHER EDUCATION PRESS

高等学校教材

统 计 学

管于华 主编

高等教育出版社

内容提要

本教材广泛吸收国内外优秀统计学教材的成果,按描述统计和推断统计搭建教材的基本框架。注重对统计方法思想的阐述,强调正确应用统计方法的条件,并配以内容丰富的实例,系统实用,脉络清晰。本教材全面介绍了统计学的基本理论和基本方法及其在经济管理中的应用,以四个大型数据集作为演算的背景数据,用 Excel 为工具完成本教材的数据处理;采用大量 Internet 链接,延伸课本内容,调动学生探求新知识的潜能。本教材还配有一定数量的习题;帮助学生巩固所学的统计概念。

本教材适用于高等学校经济、管理类本科专业;若作为专科教材,可适当删减部分章节;本教材也可供统计工作者参考。

图书在版编目(CIP)数据

统计学 / 管于华主编. —北京:高等教育出版社,
2005.8

ISBN 7-04-017720-X

I. 统... II. 管... III. 统计学-高等学校-
教材 IV. C8

中国版本图书馆 CIP 数据核字(2005)第 082100 号

策划编辑 马 丽 责任编辑 崔梅萍 封面设计 王凌波 责任绘图 尹 莉
版式设计 马静如 责任校对 殷 然 责任印制 宋克学

出版发行	高等教育出版社	购书热线	010-58581118
社 址	北京市西城区德外大街 4 号	免费咨询	800-810-0598
邮政编码	100011	网 址	http://www.hep.edu.cn
总 机	010-58581000		http://www.hep.com.cn
经 销	北京蓝色畅想图书发行有限公司	网上订购	http://www.landaco.com
印 刷	北京中科印刷有限公司		http://www.landaco.com.cn
开 本	787×960 1/16	版 次	2005 年 8 月第 1 版
印 张	23.75	印 次	2005 年 8 月第 1 次印刷
字 数	440 000	定 价	27.30 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 17720-00

前 言

随着信息技术的不断发展,无论是自然科学和社会科学的研究领域,还是国家调控和企业运作的管理过程,都会产生巨大的信息流量,人们对信息资源的掌握与利用已经成为在激烈竞争状态下取得成功的前提条件。作为信息技术支柱的统计方法,有着越来越广泛的应用前景,越来越受到人们的重视。

本教材的基本内容能够在 50~60 个课堂教学课时内完成。为了能够将统计学的基本思想和基本方法以学生乐于接受的方式传授给他们,我们在课程内容上精选、在表达方式上创新,力争使学生形成积极获取信息的意识,并掌握处理数据的技术,以适应时代发展的需要。考虑到课时的限制,我们设计了两套教学方案,针对经济、管理两个不同的方向,做一些合理的取舍。对经济类专业的学生,强调模型的建立与计算,侧重于宏观经济分析;对管理类专业的学生,强调统计在企业中的应用,如将统计过程控制等企业管理中的实用技术作为必修内容。

经济类专业		管理类专业	
课程内容	教学时间	课程内容	教学时间
第一章 总论	2	第一章 总论	2
第二章 统计学的基本问题	4	第二章 统计学的基本问题	4
第三章 统计资料的收集与整理	6	第三章 统计资料的收集与整理	6
第四章 数据的描述性分析	4	第四章 数据的描述性分析	4
第五章 指数	4	第五章 指数	4
第六章 参数估计	6	第六章 参数估计	6
第七章 假设检验	4	第七章 假设检验	4
第八章 非参数检验		第八章 非参数检验	4
第九章 方差分析		第九章 方差分析	2
第十章 相关与回归分析	6	第十章 相关与回归分析(部分)	4
第十一章 多元统计分析	4	第十一章 多元统计分析	
第十二章 时间序列分析	6	第十二章 时间序列分析	6
第十三章 统计决策	4	第十三章 统计决策	
第十四章 统计质量控制		第十四章 统计质量控制	4
合 计	50	合 计	50

本教材由南京财经大学统计学系从事本科统计学教学的全体教师集体讨论、编著完成,编著团队每个成员都热爱统计学教育事业,有丰富的教学经验,并且有良好的合作基础,2002年统计学课程被评为江苏省普通高等学校一类优秀课程。欢迎访问南京财经大学统计学系主页:<http://tjx.njue.edu.cn/>。

本教材编写分工如下:第一章,陶用之;第二章,柳杰民;第三章,张彦伟、李昌峰;第四章,朱龙杰;第五章,管于华;第六章,朱龙杰;第七章,尚卫平;第八章,管于华;第九章,沈诗霞;第十章,沈诗霞;第十一章,王芳;第十二章,王芳;第十三章,白先春;第十四章,管于华。最后由管于华总纂定稿。杨益民教授仔细阅读了本教材,并提出了许多宝贵建议。在教材编写过程中,得到院、校领导的热情支持,谨此致谢。

虽然我们竭尽全力,但由于水平有限,错误在所难免。请各位专家、读者不吝赐教,我们不胜感激。联系方式:e-mail:gggyyyhhh@hotmail.com.

获取本教材相关的教学资源有两个途径:一是登录南京财经大学网络教学平台:<http://211.87.74.54/>进入“统计学—新教材”;二是登录<http://www.njcdtjx.com/>进入“统计学—立体教材”。

编者

2005年2月

目 录

第一章 总论	1
第一节 什么是统计学	2
一、统计的涵义	2
二、统计学的过去与现在	5
三、统计学的研究方法	8
四、统计学的内容	10
第二节 统计学有什么用	11
一、统计给我们智慧与力量	11
二、统计学在经济领域的应用	12
三、统计学在管理领域的应用	13
习题	14
第二章 统计学的基本问题	17
第一节 四个数据集及数据类型	18
一、四个常用数据集	18
二、数据类型 I	20
三、数据类型 II	22
第二节 统计总体、个体与样本	23
一、统计总体和个体	23
二、总体的特点	23
三、样本	25
第三节 标志、指标与变量	25
一、统计标志	25
二、统计指标	26
三、变量	29
第四节 指标体系	29
一、统计指标体系	29
二、举例：工业企业统计指标体系——企业管理与信息处理指标子体系	30
三、举例：宏观统计指标体系——国民经济核算体系	31
第五节 统计计算工具	33

一、概述	33
二、Excel 实现数据处理的主要途径	33
习题	36
第三章 统计资料的收集与整理	39
第一节 统计资料的收集	40
一、统计资料及其来源	40
二、统计调查	41
第二节 统计资料的整理	46
一、统计资料的审定	46
二、统计资料的分组	46
三、次数分布	48
第三节 统计表与统计图	52
一、统计表	53
二、统计图	54
习题	59
第四章 数据的描述性分析	63
第一节 集中趋势的描述	64
一、数值平均数	64
二、位置平均数	70
第二节 离散程度的描述	76
一、离散程度的绝对指标	76
二、离散程度的相对指标	79
三、数据的标准化	80
四、是非标志标准差	81
第三节 分布的偏态与峰度	82
一、原点矩与中心矩	82
二、分布的偏态	83
三、分布的峰度	84
第四节 计算工具使用指南	85
一、计算器的使用	85
二、Excel 的描述统计分析功能	87
习题	90
第五章 指数	97
第一节 指数概念	98
一、什么是指数	99

二、为什么要把数据转换成指数	100
第二节 指数的构造方法	100
一、简单综合指数	101
二、简单平均比率指数	103
三、拉氏指数和派氏指数	104
四、加权平均比率指数	107
五、指数公式优良性测试与指数体系	108
第三节 指数的应用与调整	110
一、指数的调整作用	110
二、指数数列与基期更换	111
三、拉氏指数与固定权数	112
四、总指数与类指数	113
第四节 几种常用的经济指数	114
一、消费者价格指数(CPI)	114
二、商品零售价格指数	115
三、股票价格指数	116
习题	120
第六章 参数估计	125
第一节 抽样推断的基本概念与原理	127
一、抽样推断的特点和作用	127
二、重复抽样与不重复抽样	128
三、抽样误差与抽样平均误差	129
四、抽样推断的理论基础	131
五、参数估计的基本步骤	132
第二节 参数估计中的点估计	132
一、总体参数的点估计	132
二、点估计量的优良标准	133
第三节 参数估计中的区间估计	134
一、参数估计的精度与抽样平均误差计算	134
二、参数估计的误差范围与概率度	137
三、总体参数的区间估计	139
第四节 抽样组织方式及其参数估计	140
一、简单随机抽样	140
二、分层抽样	142
三、系统抽样	144

四、整群抽样	144
第五节 必要样本容量的确定	145
一、平均数的必要样本容量	145
二、成数的必要样本容量	146
三、影响必要样本容量的因素	147
习题	148
第七章 假设检验	153
第一节 假设检验概述	154
一、假设检验的基本思想	154
二、假设检验的步骤	155
三、两类错误和假设检验的规则	156
第二节 总体均值的检验	157
一、单个正态总体均值的检验	157
二、两个正态总体均值之差的检验	160
三、两个非正态总体均值之差的检验	162
第三节 总体成数的检验	166
一、单个总体成数的检验	166
二、两个总体成数之差的检验	167
第四节 总体方差的检验	168
一、一个正态总体方差的检验	168
二、两个正态总体方差之比的检验	169
习题	170
第八章 非参数检验	173
第一节 非参数检验概述	174
第二节 单样本非参数检验	175
一、 χ^2 检验	175
二、单样本 K-S 检验	177
三、符号检验	179
四、游程检验	180
第三节 两个样本的非参数检验	182
一、列联表与 χ^2 的独立性检验	182
二、两样本的 K-S 检验	184
习题	185
第九章 方差分析	189
第一节 方差分析概述	190

一、方差分析中的常用术语	190
二、方差分析的基本思想	191
三、方差分析的基本假定	192
第二节 单因素方差分析	192
一、单因素方差分析的数据结构	192
二、单因素方差分析的步骤	193
三、方差分析中的多重比较	198
第三节 双因素方差分析	199
一、双因素方差分析的种类	199
二、无交互作用的双因素方差分析	199
三、有交互作用的双因素方差分析	203
习题	208
第十章 相关与回归分析	211
第一节 相关分析	212
一、相关关系的概念及分类	212
二、相关关系的识别	214
第二节 等级相关分析	219
一、Spearman 等级相关系数	219
二、Kendall 等级相关系数	221
第三节 一元线性回归分析	223
一、一元线性回归	223
二、参数 β_0, β_1 的最小二乘估计	224
三、对一元回归方程的评价	226
四、一元回归方程的预测区间	230
第四节 多元线性回归分析	233
一、多元线性回归模型	233
二、多元回归模型的参数估计	234
三、对多元线性回归方程的评价	235
第五节 非线性模型	238
一、直接代换法	238
二、间接代换法	239
习题	239
第十一章 多元统计分析	245
第一节 主成分分析	246
一、基本思想	246

二、数学模型	247
三、模型的求解	248
四、主成分的性质	249
五、基本步骤与应用实例	249
第二节 因子分析	254
一、基本思想	254
二、数学模型	254
三、因子载荷的统计含义	255
四、因子的求解	256
五、因子得分	256
六、基本步骤与应用实例	256
第三节 聚类分析	258
一、基本思想	259
二、统计量	259
三、分类方法(系统聚类法)	260
四、基本步骤与应用实例	261
第四节 判别分析	266
一、基本思想	266
二、基本方法	266
三、判别效果的评价	267
四、基本步骤与应用实例	268
习题	270
第十二章 时间序列分析	273
第一节 时间序列及其因素分解	274
一、时间序列的类型	275
二、构成因素	275
三、时间序列的数学模型	276
第二节 时间序列的描述性分析	276
一、图形描述	276
二、指标分析	278
第三节 因素分析	287
一、长期趋势分析	287
二、季节变动分析	291
三、循环波动分析	294
习题	297

第十三章 统计决策	303
第一节 统计决策概述	304
一、决策的概念	304
二、决策的步骤	305
第二节 风险型决策方法	306
一、损益矩阵	306
二、以期望值为标准的决策方法	307
三、以合理性为标准的决策方法	308
四、以最大可能性为标准的决策方法	308
五、决策树	309
六、敏感性分析	310
第三节 贝叶斯决策方法	312
一、贝叶斯定理	312
二、贝叶斯决策分析	313
第四节 不确定型决策方法	316
一、“好中求好”决策方法	317
二、“坏中求好”决策方法	318
三、 α 系数决策方法	319
四、“最小的最大后悔值”决策方法	320
五、各种决策方法的比较	320
习题	321
第十四章 统计质量管理	325
第一节 质量与质量管理	326
一、质量与质量波动原因	326
二、质量管理与统计质量管理	327
三、质量管理的简要历程	327
第二节 质量管理中的两个常用图形	328
一、帕累托图	328
二、鱼刺图	330
第三节 统计过程控制图	331
一、计量值控制图	332
二、计数值控制图	336
三、控制图的观察分析	337
第四节 工序能力分析	340
一、工序能力指数	341

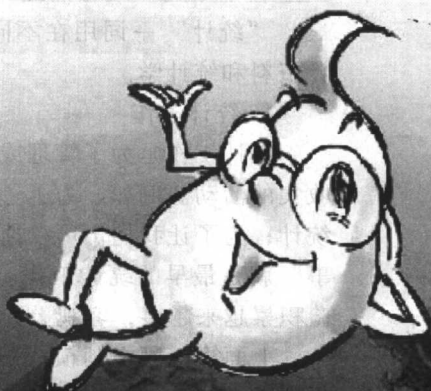
二、工序能力评价	343
三、控制图	344
习题	344
附表一 标准正态分布表	348
附表二 随机数表	351
附表三 t 分布临界值表	352
附表四 χ^2 分布临界值表	354
附表五 F 分布临界值表	355
附表六 单样本 K-S 检验统计量表	357
附表七 符号检验界域表	358
附表八 游程检验临界值表	359
附表九 相关系数临界值表	360
附表十 Spearman 等级相关系数临界值表	361
附表十一 Kendall 等级相关系数临界值表	362
附表十二 控制图系数表	363
参考文献	364
后记	365

第一章 总 论

通过本章的学习，我们应该知道：

1. 什么是统计学？
2. 统计学有什么用？

统计能把数据变为有用的信息
统计能给我们智慧与力量



学生：统计学？估计是和数据有关。我这见数字就头痛的人，学这门课可能有难度。要说花工夫，我倒不在乎，可学了有什么用？值得花大功夫吗？数海茫茫，真有渡水之舟？

第一节 什么是统计学

一、统计的涵义

在人们的一般认识中，“统计”就是“计数”。小至一个人、家庭，大至一个企业、国家都有计数的任务。一个月的收入、一年的利润都会是我们经常关心盘算的，这些数据就是统计的成果。世界各国都有各自的官方统计部门，负责对人口、资源、环境和社会经济活动等各方面进行“计数”，并将这些数据资料以公共产品的方式定期公布，往往命名为“统计年鉴”。现在关心这些数据的不仅仅是政治家、经济学家和各方学者，普通老百姓由于其收入可能和物价水平相联系，所以他们可能不知道CPI(消费者价格指数)的确切含义，但对物价指数还是会投以关注。

网络链接：中华人民共和国国家统计局 <http://www.stats.gov.cn/>

“统计”一词用在不同的场合，可以替代的涵义有三个：统计工作、统计资料和统计学。

1. 统计工作

统计工作是为了管理或认识的需要，对社会经济现象和自然现象进行数据收集的活动。例如，为了了解国家的家底，我们进行人口、自然资源和财富的统计；为了让我们仍可以拥有蓝天白云，我们进行环境质量的统计。“结绳记事”就是最早的统计工作。统计工作的延续与发展使得其范围越来越大、经验积累越来越多，客观上为统计学的产生和发展奠定了基础。

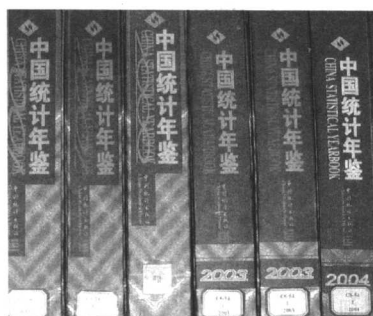
上面第一段中画横线的那句话，严格来说应该是“这些数据就是统计工作的成果”。

2. 统计资料

统计资料是统计工作过程中所取得的各项数字资料以及与之相关信息的总称。统计资料是统计工作取得的能够说明所研究对象的数据。例如国内生产总值数据,说明整个国家的生产规模;居民消费支出,说明居民的生活水平。这些数据经常会在报纸、杂志上出现。随着信息技术的发展与网络的普及,统计资料的公布不再仅仅是纸质资料了,大量的电子版的数据可以方便地从各国官方统计网站上获得,大部分都是免费的。在国外的统计局网站中,专家评价最高的是荷兰统计局的网站,网页显著位置的“Statline”是荷兰统计局的电子数据库,通过它可以查找各类统计数据,并且可以免费下载。

网络链接:荷兰国家统计局 <http://www.cbs.nl/en/>

我国统计资料的发布途径越来越规范。官方的统计数据通过“中国统计年鉴”和各省、市、地区的统计年鉴定期出版,一般都会同时提供纸质和电子两种形式。目前各级统计局都建立了自己的网站,及时提供有关数据,比如,国家统计局就以“进度统计数据”、“年度统计数据”(见图 1.1)、“普查数据”、“专题数据”和“部门数据”五个栏目及时公布许多重要指标。



年度统计数据			
综合	~年度~	📄	📄 检索
国民经济核算	~年度~	📄	📄 检索
人口	~年度~	📄	📄 检索
劳动工资与就业	~年度~	📄	📄 检索
固定资产投资	~年度~	📄	📄 检索
能源生产和消费	~年度~	📄	📄 检索
财政	~年度~	📄	📄 检索
物价	~年度~	📄	📄 检索
农村住户	~年度~	📄	📄 检索
城镇住户	~年度~	📄	📄 检索
农业	~年度~	📄	📄 检索
工业	~年度~	📄	📄 检索
建筑业	~年度~	📄	📄 检索
运输和邮电	~年度~	📄	📄 检索
国内贸易	~年度~	📄	📄 检索
社会服务业	~年度~	📄	📄 检索
旅游	~年度~	📄	📄 检索
金融和保险	~年度~	📄	📄 检索
教育科技和文化	~年度~	📄	📄 检索
体育卫生和社会福利	~年度~	📄	📄 检索
香港特别行政区主要社会经济指标	~年度~	📄	📄 检索
澳门特别行政区主要社会经济指标	~年度~	📄	📄 检索
台湾省主要社会经济指标	~年度~	📄	📄 检索

图 1.1 中国统计资料的获取渠道

教师:作为一个普通公民,我们有获取统计资料的权利。在知识经济时代,无论是有造诣的学者还是成功的企业家都不能漠视这一巨大的“财富”宝库。“数据就是金钱”千真万确。

3. 统计学

统计学是在统计工作的经验积累到一定程度时自然产生的,它是收集、整理、描述分析统计数据的方法和技术,为我们的决策提供“量”方面的依据。《不列颠百科全书》对统计学的定义为“收集和分析数据的科学和艺术(art)”。

这里提到的“艺术”并不是说统计学属于音乐、美术一类的艺术范畴，而是指统计方法的应用具有技巧性、技术性，它展现数据时往往让我们能享受美感和实用。也许你会感叹埃及的金字塔，可你知道人口统计学中也有一个“金字塔”吗？它同样美丽，见图 1.2。

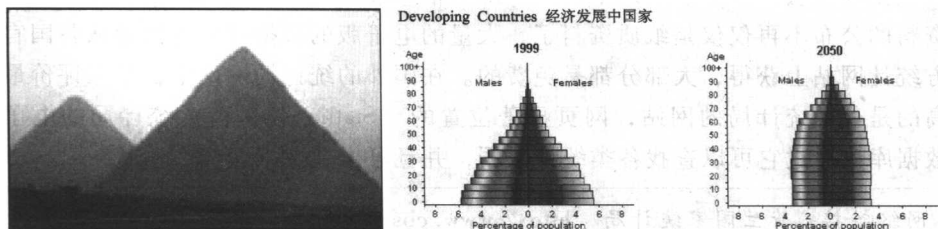


图 1.2 埃及金字塔与“人口金字塔”

统计学的定义突出了统计学研究对象的两个方面：第一，收集数据；第二，分析数据。收集数据的目的是为了解决某一应用或理论上的问题。但是单有一堆杂乱无章的数据，什么问题也回答不了，用处不大，需要经过整理，从中发掘有用的信息并用适当的形式表达出来，然后用科学的方法进行分析，以针对所研究的问题做出一定的结论。下面我们举一个最简单的例子。

随着互联网使用的日益普及，人们对相关情况也越来越关注。互联网的用户人数及分布、信息流量分布、域名注册等信息以及用户每月实际花费的上网费用、用户平均每周上网时间、用户通常在什么时间上网等资料都可以从电信局或网络公司的业务记录中获得，但众多的信息资料往往会让我们觉得茫然，所以进一步数据整理与信息展现是必不可少的。我们可以汇总每天某个时间点上，在线用户占注册用户的比例来反映用户通常上网的时间，见表 1.1。

表 1.1 用户通常上网时间

时间	上网比例	时间	上网比例	时间	上网比例
0 点	14.5%	8 点	11.8%	16 点	31.2%
1 点	5.6%	9 点	22.5%	17 点	27.7%
2 点	4.2%	10 点	24.6%	18 点	22.1%
3 点	2.6%	11 点	22.6%	19 点	31.3%
4 点	2.5%	12 点	22.6%	20 点	44.0%
5 点	1.6%	13 点	24.3%	21 点	43.9%
6 点	1.7%	14 点	28.6%	22 点	35.0%
7 点	2.7%	15 点	31.4%	23 点	22.2%

资料来源：中国互联网络信息中心 中国互联网络发展状况统计报告(2003)