

增刊



21世纪

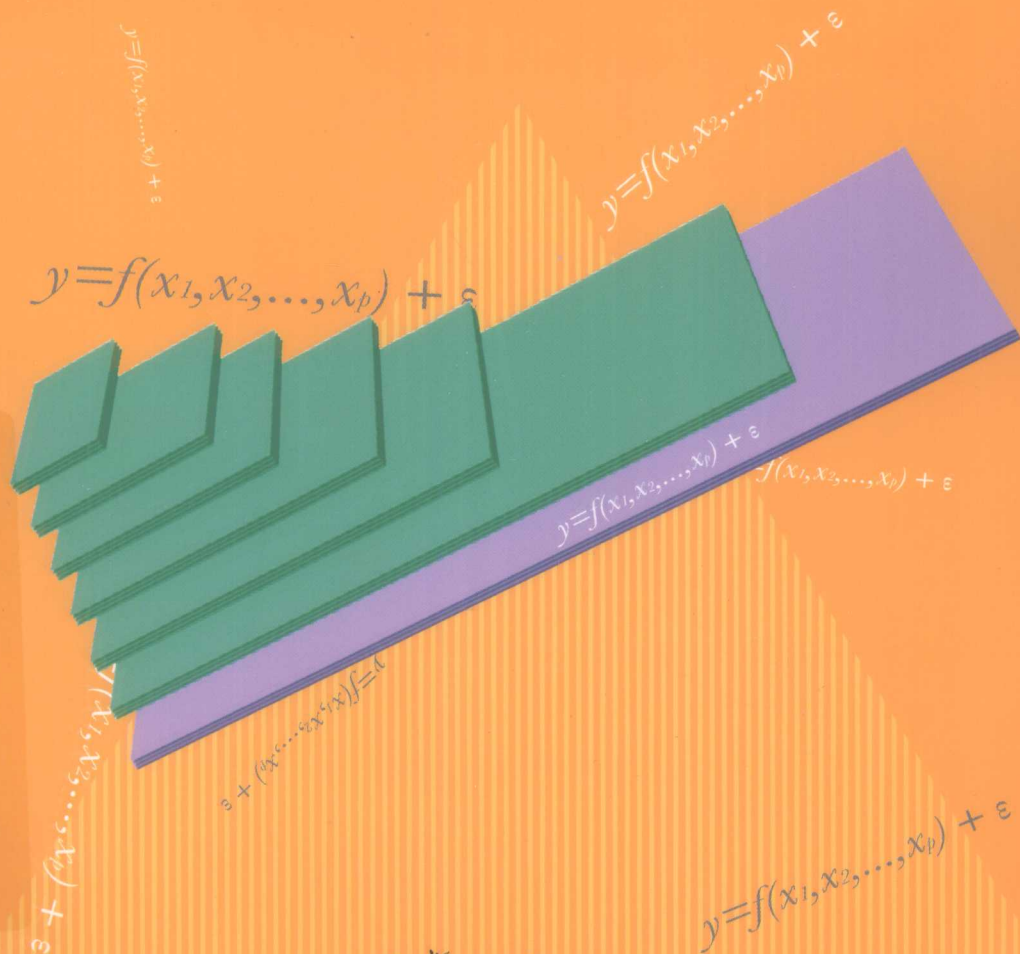
顾问 王静龙 艾春荣 徐国祥 周 勇

高校统计学专业教材系列

上海市教委重点课程建设项目
上海财经大学精品课程

应用时间序列分析

王黎明 王 连 杨 楠 编著



复旦大学出版社

顾问 王静龙 艾春荣 徐国祥 周勇

21世纪



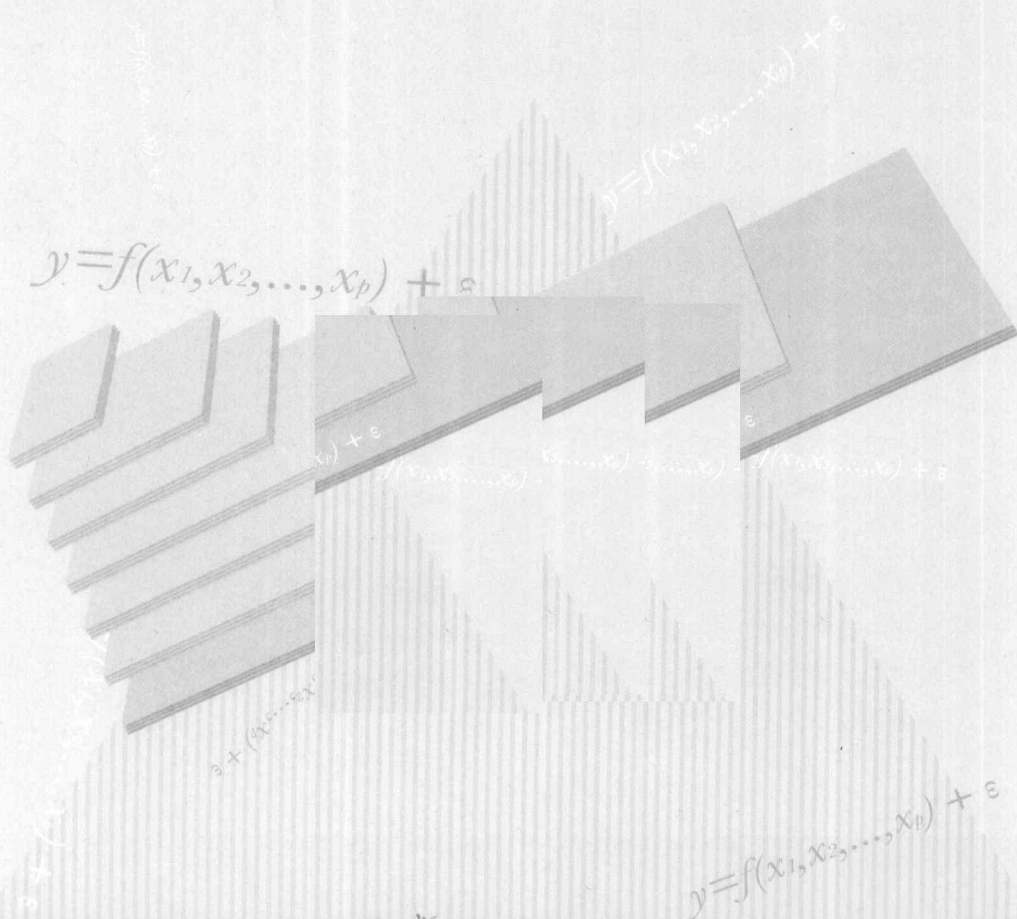
21世纪

上海市教委重点课程建设项目
上海财经大学精品课程

高校统计学专业教材系列

应用时间序列分析

王黎明 王连 杨楠 编著



复旦大学出版社

图书在版编目(CIP)数据

应用时间序列分析/王黎明,王连,杨楠编著. —上海:复旦大学出版社,2009.9
(复学博学·21世纪高校统计学专业教材系列)
上海市教委重点课程建设项目
ISBN 978-7-309-06880-1

I. 应… II. ①王…②王…③杨… III. 时间序列分析-高等学校-教材
IV. 0211.61

中国版本图书馆CIP数据核字(2009)第165420号

应用时间序列分析

王黎明 王 连 杨 楠 编著

出版发行 复旦大学出版社 上海市国权路579号 邮编 200433
86-21-65642857(门市零售)
86-21-65100562(团体订购) 86-21-65109143(外埠邮购)
fupnet@fudanpress.com <http://www.fudanpress.com>

责任编辑 王联合

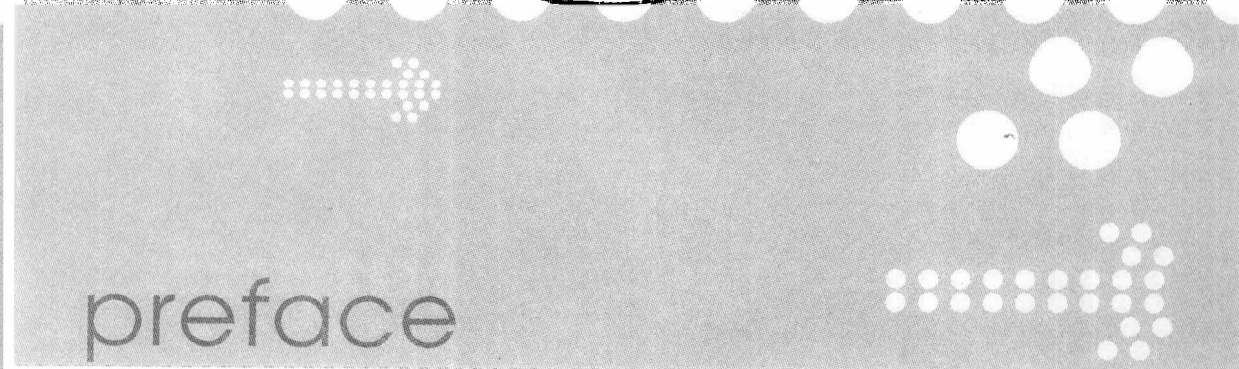
出品人 贺圣遂

印 刷 上海肖华印务有限公司
开 本 787×960 1/16
印 张 18.75
字 数 357千
版 次 2009年9月第一版第一次印刷

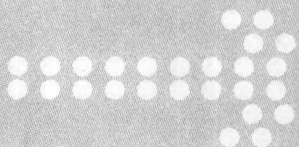
书 号 ISBN 978-7-309-06880-1/O·434
定 价 32.00元

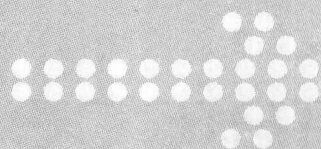
如有印装质量问题,请向复旦大学出版社发行部调换。

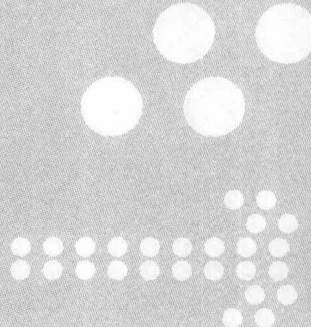
版权所有 侵权必究



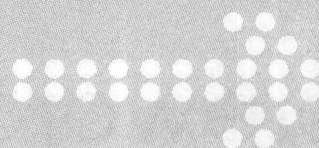
preface

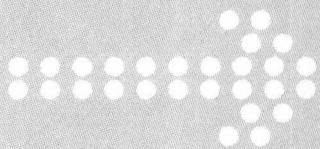






foreword





本书是我们多年教学和科研工作的积累,2001年以来,本书的作者一直为上海财经大学统计学专业的本科生和硕士研究生讲授时间序列分析课程,书中的大部分内容为多年来不断改进的讲义,其中部分案例为体现其典型性也引用了他人著作。在此,我们谨向对本书出版给予帮助的同行人和朋友表示衷心的感谢。全部工作的完成也是我们多年友好合作的结果,研究生王帅同学参加了部分章节和习题的编写和整理工作,撰写了大部分的 Eviews 软件实现的案例分析,同时也参加了最后的统稿和校对工作。

为了便于教师上课,我们将全书所用的数据和 PowerPoint 课件刻录成光盘附赠给授课老师,需要者可直接发邮件至 unionw@sina.com 索取。

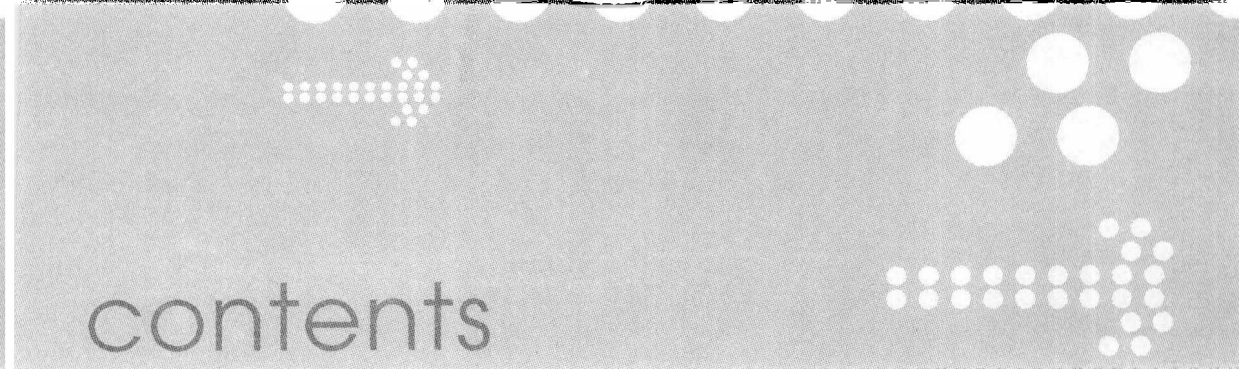
由于编者的水平有限,在取材及其结构上,难免会存在不够妥当的地方,错误之处也在所难免,恳请同行专家和广大读者能给我们宝贵的批评和建议。

编者

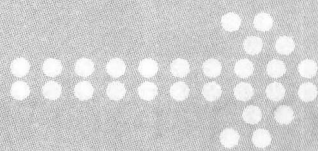
王黎明 王 连 杨 楠

2009 年 3 月

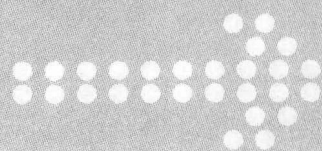
于上海财经大学凤凰楼

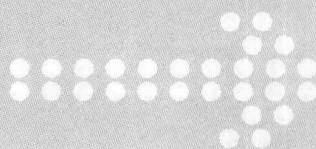


contents



第四章 非平稳时间序列和季节序列模型	81
§ 4.1 均值非平稳	82
§ 4.2 自回归求和移动平均模型 (ARIMA)	84
§ 4.3 方差和自协方差非平稳	87
§ 4.4 季节时间序列 (SARIMA) 模型	89
习题四	90
第五章 时间序列的模型识别	92
§ 5.1 自相关和偏自相关系数法	93
§ 5.2 F 检验法	99
§ 5.3 信息准则法	101
习题五	107
第六章 时间序列模型参数的统计推断	109
§ 6.1 自协方差系数的参数估计	109
§ 6.2 $ARMA(p, q)$ 模型参数的矩估计	112
§ 6.3 $ARMA(p, q)$ 模型参数的极大似然估计	118
§ 6.4 $ARMA(p, q)$ 模型参数的最小二乘估计	127
§ 6.5 $ARMA(p, q)$ 模型的诊断检验	128
§ 6.6 $ARMA(p, q)$ 模型的优化	130
习题六	132
EViews 软件介绍 (Ⅲ)	133
第七章 平稳时间序列模型预测	140
§ 7.1 最小均方误差预测	140
§ 7.2 对 AR 模型的预测	143
§ 7.3 MA 模型的预测	146





第十一章 (超)高频数据的建模与分析简介	261
§ 11.1 (超)高频数据的特点	262
§ 11.2 (超)高频数据与 ACD 模型	263
§ 11.3 交易持续期的集聚性	267
§ 11.4 UHF-GARCH 模型	268
习题十一	269
 附录 1 数据	 271
 附录 2 常见分布表	 280
 参考文献	 286

第一章

时间序列分析概论

人类为了探索周围的世界,常常依时间发展的先后顺序进行观测。这些观测到的数据有着比较独特的性质,即将来的数据通常以某种随机的方式依赖于当前得到的观测数据,而这种相依性使得利用过去预测未来成为可能。时间序列是变量按时间间隔的顺序而形成的随机变量序列。大量自然界、社会经济等领域的统计指标都依年、季、月或日统计其指标值,随着时间的推移,形成了统计指标的时间序列。因此,时间序列是某一统计指标长期变动的数量表现。时间序列分析就是估算和研究某一时间序列在长期变动过程中所存在的统计规律性。本章将介绍时间序列分析的基本思想和一般理论。

§ 1.1 时间序列的定义和例子

在统计学研究中,有大量的数据是按照时间顺序排列的,用数学方法来表述就是使用一组随机序列

$$\cdots, X_1, X_2, \cdots, X_t, \cdots \quad (1.1)$$

表示随机事件的时间序列,简记为 $\{X_t, t \in T\}$ 或者 $\{X_t\}$ 。

类似于样本与样本观测值的关系,我们可以使用

$$x_1, x_2, \cdots, x_n \quad (1.2)$$

表示上述时间序列(1.1)的 n 个有序观测值,称其为序列长度为 n 的观测值序列。

在时间序列问题中,数据的时间顺序是重要的,时间序列的一个显著的特征就是记录的相依性。一般来说,关于时间序列 $\{X_t\}$,对于任意的 t , X_t 是一个随机变量,且每个随机变量所服从的分布可以不同,对于任意的 t 和 s , X_t 与 X_s 不是相互独立的。时间序列的应用背景十分广泛,依照不同的需要,数据的收集可以按小时、天、周、月或者年为间隔进行,现在更有以秒为时间间隔的所谓高频时间序列。下面我们介绍一些不同领域中的实际数据例子,相应的原始数据在书后以附录形式给出。

例 1.1 1820—1869 年的太阳黑子数据(单位:个):太阳黑子是太阳表面上的黑点,它反映了太阳振动的全部演变,与太阳发生电效应的行为有关。1820—1869 年的太阳黑子数依时间画在图 1.1 中。该图中,横轴是时间指标 t (在这里的 t 以年为单位),纵轴表示在时间 t 内太阳黑子个数的观测值 X_t ,这种图称为时间序列图,简称时序图。

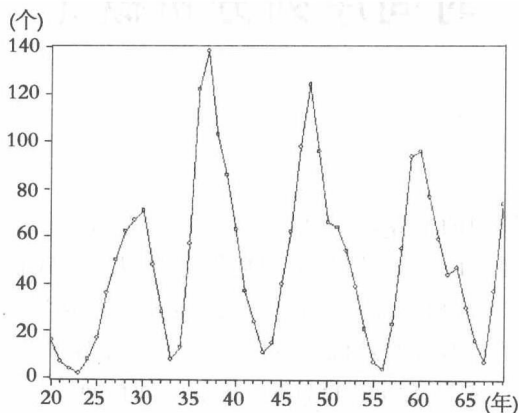


图 1.1 1820—1869 年间太阳黑子数据

例 1.2 1985—2007 年我国居民消费价格指数(CPI):居民消费价格指数,即消费者物价指数(Consumer Price Index),英文缩写 CPI,是反映与居民生活有关的产品及劳务价格统计出来的物价变动指标,通常作为观察通货膨胀水平的重要指标。如果消费者物价指数升幅过大,表明通胀已经成为经济不稳定因素。本例给出了我国 1985—2007 年的 CPI 年度数据时间序列图,如图 1.2 所示。

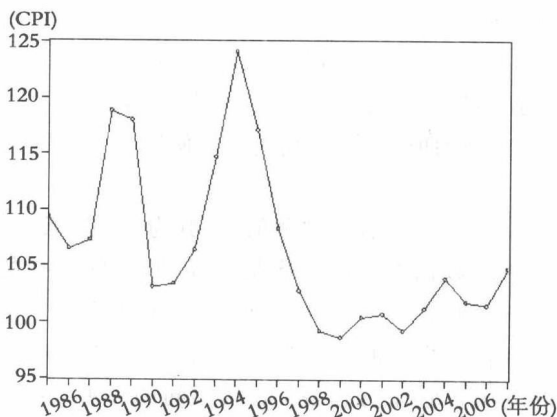


图 1.2 1985—2007 年我国居民消费价格指数

例 1.3 1978—2007 年我国 GDP 数据(单位:亿元):GDP 即国内生产总

值,它是对一国(地区)经济在核算期内所有常驻单位生产的最终产品总量的度量,常常被看成反映一个国家(地区)经济状况的重要指标。本例给出我国 1978—2007 年 GDP 数据(单位:亿元)的时间序列图,见图 1.3。

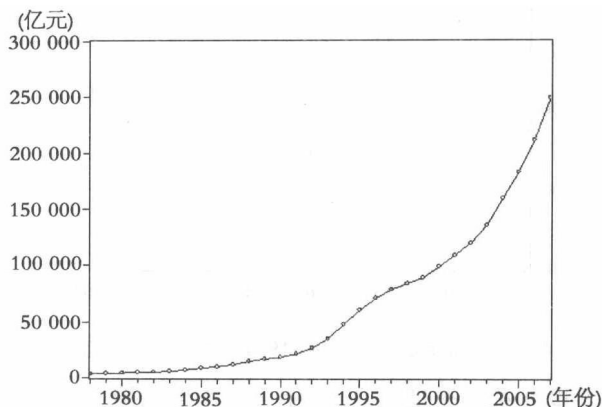


图 1.3 1978—2007 年我国 GDP

例 1.4 北京在历史上也是自然灾害频发的地区,在各种自然灾害中,水旱灾害发生的次数最多,危害最大。表 1.1 列出了北京地区 1949—1964 年的洪涝灾害面积数据(单位:万亩)。

表 1.1 北京地区 1949—1964 年的洪涝灾害面积

年份	受灾面积	年份	受灾面积
1949	331.12	1957	25.00
1950	380.44	1958	84.72
1951	59.63	1959	260.89
1952	37.89	1960	27.18
1953	103.66	1961	20.74
1954	316.67	1962	52.99
1955	208.72	1963	99.25
1956	288.78	1964	55.36

我们使用 X_1 表示第一年(1949 年)的受灾面积, X_2 表示第二年(1950 年)的受灾面积…… $X_1, X_2, \dots$ 是一列按照时间顺序排列的随机序列,所以是时间序列。 $x_1, x_2, \dots, x_{16}$ 是北京地区 1949—1964 年的洪涝受灾面积,则

$$x_1 = 331.12, x_2 = 380.44, \dots, x_{16} = 55.36$$

是时间序列 $\{X_t\}$ 的样本观测值,样本容量为 16,它是时间序列 $\{X_t\}$ 的一次实现的一部分。时间序列 $\{X_t\}$ 的样本观测值 $x_1, x_2, \dots, x_{16}$ 可以由图 1.4 表示。

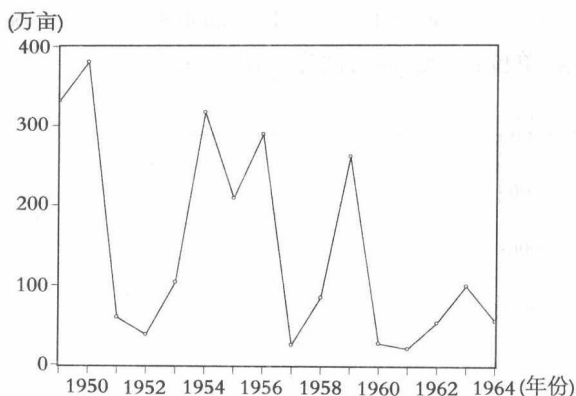


图 1.4 北京地区 1949—1964 年的洪涝灾害面积

例 1.5 1992 年第一季度至 2008 年第三季度我国 GDP 季度数据(单位:亿元):图 1.5 是这一季度指标的时间序列图。

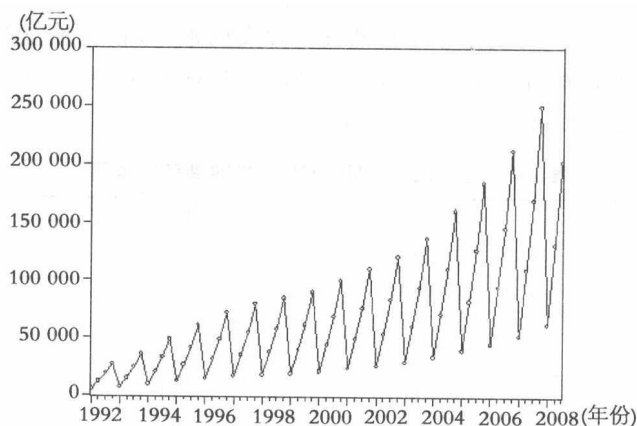


图 1.5 1992 年第一季度至 2008 年第三季度我国 GDP

例 1.6 1997 年 1 月—2008 年 9 月美元对人民币汇率的月度数据(见图 1.6,单位:元):2005 年 7 月 21 日中国启动人民币汇率改革以来,不断完善汇率形成机制,人民币对美元汇率总体呈现小幅上扬态势。过去两年多,人民币累计升值近 16%。2008 年以来,人民币汇率升幅已接近 4.5%。

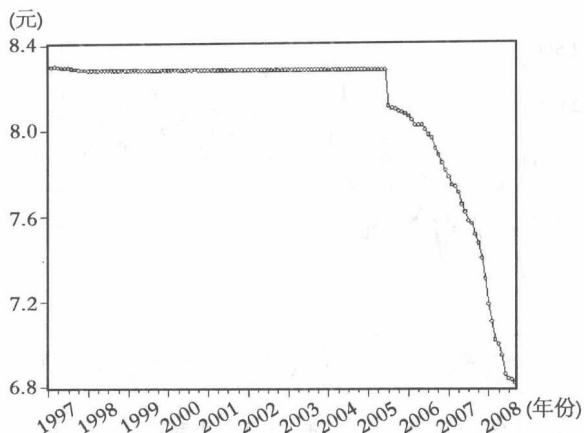


图 1.6 1997 年 1 月—2008 年 9 月美元对人民币汇率

例 1.7 1990 年 12 月 19 日—2008 年 11 月 6 日上证 A 股指数日数据(除去节假日,共 4 386 个数据,见图 1.7):上证 A 股指数的样本股为 A 股,自 1990 年 12 月 19 日正式发布。而上证综合指数的样本股是全部上市股票,包括 A 股和 B 股。上证综合指数从总体上反映了上海证券交易所上市股票价格的变动情况,自 1991 年 7 月 15 日起正式发布。

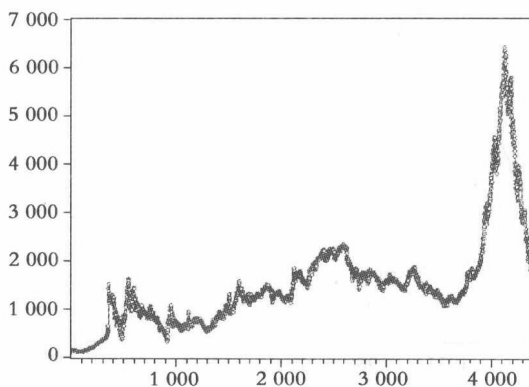


图 1.7 1990 年 12 月 19 日—2008 年 11 月 6 日上证指数

例 1.8 1980 年 1 月—1991 年 10 月澳大利亚红酒的月度销量(单位:立方分米),共包括 142 个数据,见图 1.8。