

大数据时代经济与金融数据分析系列丛书



# 经济金融计量及其R语言应用

朱顺泉 编著



清华大学出版社

大数据时代经济与金融数据分析系列丛书



# 经济金融计量及其R语言应用

朱顺泉 编著

清华大学出版社  
北京

## 内 容 简 介

本书结合大量精选的实例全面介绍使用 R 语言进行经济与金融分析的方法。全书共 14 章,内容包括经济与金融计量学绪论,R 语言的下载、安装与启动,R 语言对象与数据存取,参数估计与假设检验的 R 语言应用,线性回归分析的 R 语言应用,多重共线性的 R 语言应用,异方差问题的 R 语言应用,自相关问题的 R 语言应用,时间序列分析 ARIMA 模型预测的 R 语言应用,单位根、协整与格兰杰因果检验的 R 语言应用,时间序列分析 GARCH 模型的 R 语言应用,面板数据分析的 R 语言应用,基于 R 语言的金融数据分析综合应用,创业板科技型上市公司股权激励对其价值影响的计量检验研究。

本书内容新颖、全面,实用性强,融理论、方法、应用于一体,可供统计学、数量经济学、管理科学与工程、应用数学、计算数学、概率统计、金融学、金融工程、投资学、金融专业硕士、经济学、财务管理、会计学、工商管理等专业的本科高年级学生与研究生使用。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

### 图书在版编目(CIP)数据

经济金融计量及其 R 语言应用/朱顺泉编著. —北京:清华大学出版社,2016

大数据时代经济与金融数据分析系列丛书

ISBN 978-7-302-43795-6

I. ①经… II. ①朱… III. ①程序语言—程序设计—应用—金融—计量经济学 IV. ①F83

中国版本图书馆 CIP 数据核字(2016)第 100137 号

责任编辑:刘向威 战晓雷

封面设计:文 静

责任校对:梁 毅

责任印制:沈 露



出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, [c-service@tup.tsinghua.edu.cn](mailto:c-service@tup.tsinghua.edu.cn)

质量反馈:010-62772015, [zhiliang@tup.tsinghua.edu.cn](mailto:zhiliang@tup.tsinghua.edu.cn)

课 件 下 载: <http://www.tup.com.cn>, 010-62795954

印 装 者:三河市少明印务有限公司

经 销:全国新华书店

开 本:185mm×260mm 印 张:14 字 数:348 千字

版 次:2016 年 8 月第 1 版 印 次:2016 年 8 月第 1 次印刷

印 数:1~2000

定 价:30.00 元



大数据时代,数据成为商务决策最为重要的参考之一,数据分析行业迈入了一个全新的阶段。R语言是一款非常优秀的数据统计分析与图形展示软件,本书侧重于使用R语言进行经济与金融数据分析,同时结合大量精选的实例问题对R语言进行科学、准确和全面的介绍,以使读者能深刻理解R语言的精髓和灵活、高效的使用技巧。

R语言是经济金融数据分析、微宏观经济预测等领域应用非常广泛的计量软件之一,它具有免费、统计与作图功能强、数据接口丰富、短小精悍、运算速度快、更新和发展速度惊人等特点,因而受到广大用户的欢迎和喜爱。本书通过丰富的经济与金融实例,详细介绍了R3.1.2在经济与金融领域中的应用,书中所有运算都在Ri386 V3.1.2上调试通过。

本书侧重于理论与应用相结合,实例丰富且通俗易懂,重点讨论了R语言与经济与金融计量分析应用等,详细介绍了各种计量方法在R语言中的应用过程。本书适合作为经济学、金融学、统计学等相关专业的本科生或研究生学习经济计量学、金融计量学等课程的实验参考用书,同时对从事经济与金融数据分析的实际工作者也大有裨益。通过本书,读者不仅能掌握R语言及其程序包本身的应用,而且能学会从实际问题分析入手,利用R语言进行经济与金融计量分析,并对结果进行分析。

本书实例与内容丰富,有很强的针对性,书中各章详细地介绍了实例的R具体操作过程,读者只需按照书中介绍的步骤一步一步地实际操作,就能掌握全书的内容。

本书的内容安排如下:第1章为经济与金融计量学绪论;第2章介绍R语言的下载、安装与启动;第3章介绍R语言对象与数据存取;第4章介绍参数估计与假设检验的R语言应用;第5章介绍线性回归分析的R语言应用;第6章介绍多重共线性的R语言应用;第7章介绍异方差问题的R语言应用;第8章介绍自相关问题的R语言应用;第9章介绍时间序列分析ARIMA模型预测的R语言应用;第10章介绍单位根、协整与格兰杰因果检验的R语言应用;第11章介绍时间序列分析GARCH模型的R语言应用;第12章介绍面板数据分析的R语言应用;第13章是基于R语言的金融数据分析综合应用;第14章是创业板科技型上市公司股权激励对其价值影响的计算检验研究(广东省科技计划项目软科学阶段性成果。编号:2015A070704058)。

本书的出版得到了清华大学出版社的支持和帮助,应该感谢他们为读者提供了一个好的平台!由于时间和水平的限制,书中难免出现一些纰漏,恳请读者提出宝贵意见。

作者

2016年3月于广州



<b>第 1 章 经济与金融计量学绪论</b> .....	1
1.1 经济计量学与金融计量学的含义及建模步骤 .....	1
1.1.1 计量经济学与金融计量经济学的含义.....	1
1.1.2 经济计量学与金融计量学建模过程.....	1
1.1.3 经济与金融模型中的数据.....	3
1.2 经济与金融计量软件简介 .....	4
1.2.1 R 软件简介.....	4
1.2.2 Python 软件简介 .....	4
1.2.3 Stata 软件简介 .....	4
1.2.4 EViews 软件简介 .....	5
1.2.5 SAS 软件简介 .....	5
1.2.6 Matlab 软件简介 .....	5
1.2.7 SPSS 软件简介 .....	5
练习题.....	6
<b>第 2 章 R 语言的下载、安装与启动</b> .....	7
2.1 选择 R 语言的理由 .....	7
2.2 R 语言下载 .....	8
2.3 R 语言安装 .....	9
2.4 R 语言程序包的安装.....	10
2.5 R 语言的启动.....	11
2.6 R 语言的退出.....	12
2.7 R 语言的在线帮助系统.....	12
练习题 .....	13
<b>第 3 章 R 语言对象与数据存取</b> .....	14
3.1 R 语言的对象与属性.....	14
3.2 对象信息的浏览和删除.....	16
3.3 向量对象.....	17
3.3.1 数值型向量对象 .....	17
3.3.2 字符型向量对象 .....	18

3.3.3	逻辑型向量 .....	18
3.3.4	因子型向量 .....	19
3.3.5	数值型向量的运算 .....	20
3.3.6	常用统计函数 .....	21
3.3.7	向量的下标与子集(元素)的提取 .....	22
3.4	数组与矩阵对象 .....	24
3.4.1	数组的建立 .....	24
3.4.2	矩阵的建立 .....	25
3.4.3	数组与矩阵的下标与子集(元素)的提取 .....	27
3.4.4	矩阵的运算函数 .....	28
3.5	数据框对象 .....	31
3.5.1	数据框的直接建立 .....	31
3.5.2	数据框的间接建立 .....	32
3.5.3	适用于数据框的函数 .....	32
3.5.4	数据框的下标与子集的提取 .....	33
3.5.5	数据框中添加新变量 .....	34
3.6	时间序列对象 .....	35
3.7	列表对象 .....	36
3.8	R 语言数据存储 .....	37
3.9	R 语言数据读取 .....	38
3.9.1	文本文件数据的读取 .....	38
3.9.2	Excel 数据的读取 .....	40
3.9.3	R 语言中数据集的读取 .....	41
3.9.4	R 语言中的格式数据 .....	42
3.10	R 语言编程 .....	42
3.10.1	R 语言函数基础 .....	43
3.10.2	循环和向量化 .....	44
3.10.3	用 R 语言编写程序 .....	45
3.10.4	用 R 语言编写函数 .....	46
	练习题 .....	46

## 第 4 章 参数估计与假设检验的 R 语言应用 .....

4.1	参数估计的 R 语言应用 .....	48
4.1.1	点估计矩分析法的 R 语言应用 .....	48
4.1.2	单正态总体均值区间估计的 R 语言应用 .....	49
4.1.3	单正态总体方差区间估计的 R 语言应用 .....	52
4.2	假设检验的 R 语言应用 .....	53
4.2.1	参数假设检验的基本理论 .....	53
4.2.2	单个样本 $t$ 检验的 R 语言应用 .....	56

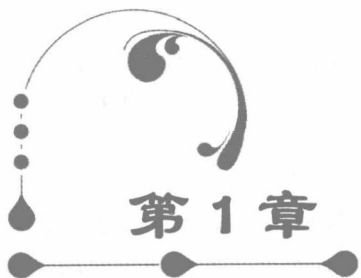
4.2.3	两个独立样本 $t$ 检验的 R 语言应用	57
4.2.4	配对样本 $t$ 检验的 R 语言应用	59
4.2.5	单样本方差假设检验的 R 语言应用	61
4.2.6	双样本方差假设检验的 R 语言应用	62
	练习题	64
<b>第 5 章</b>	<b>线性回归分析的 R 语言应用</b>	<b>66</b>
5.1	一元线性回归分析基本理论	66
5.1.1	一元线性回归分析的 OLS 估计	66
5.1.2	一元线性回归模型的统计检验	67
5.1.3	一元线性回归模型预测的置信区间	70
5.2	一元线性回归分析的 R 语言应用	70
5.3	多元线性回归分析基本理论	73
5.3.1	多元线性回归模型假设	73
5.3.2	多元线性回归模型的矩阵解法	74
5.3.3	多元线性回归模型的统计检验	74
5.4	多元线性回归分析的 R 语言应用	76
5.5	稳健线性回归分析的 R 语言应用	80
5.5.1	线性回归中的几个术语	80
5.5.2	数据描述	81
5.5.3	普通最小二乘(OLS)回归的 R 语言应用	82
5.5.4	稳健回归的 R 语言应用	84
	练习题	86
<b>第 6 章</b>	<b>多重共线性的 R 语言应用</b>	<b>93</b>
6.1	多重共线性的概念	93
6.2	多重共线性的后果	94
6.3	产生多重共线性的原因	94
6.4	多重共线性的识别和检验	95
6.5	消除多重共线性的方法	96
6.6	多重共线性诊断的 R 语言应用	98
6.7	多重共线性消除的 R 语言应用	100
	练习题	102
<b>第 7 章</b>	<b>异方差问题的 R 语言应用</b>	<b>104</b>
7.1	异方差的概念	104
7.2	异方差产生的原因	105
7.3	异方差的后果	106
7.4	异方差的识别检验	106

7.4.1	根据问题的经济背景,分析是否可能存在异方差	106
7.4.2	图示法	107
7.4.3	统计检验方法	107
7.5	消除异方差的方法	108
7.6	异方差诊断的 R 语言应用	110
7.7	异方差消除的 R 语言应用	111
	练习题	111
<b>第 8 章</b>	<b>自相关问题的 R 语言应用</b>	<b>113</b>
8.1	自相关的概念	113
8.2	产生自相关的原因	113
8.3	自相关的后果	114
8.4	自相关的识别和检验	115
8.5	自相关的处理方法	117
8.6	自相关性诊断的 R 语言应用	119
8.7	自相关消除的 R 语言应用	121
	练习题	121
<b>第 9 章</b>	<b>时间序列分析 ARIMA 模型预测的 R 语言应用</b>	<b>127</b>
9.1	ARIMA 模型	127
9.2	通过差分得到平稳时间序列	128
9.3	确定合适的 ARIMA 模型	130
9.4	ARIMA 模型预测	132
9.5	ARIMA 模型预测结果的检验	133
	练习题	134
<b>第 10 章</b>	<b>单位根、协整与格兰杰因果检验的 R 语言应用</b>	<b>135</b>
10.1	时间序列分析的基本理论	135
10.1.1	平稳、协整、因果检验的基本概念	135
10.1.2	单位根检验	136
10.1.3	协整检验	137
10.1.4	误差修正模型	138
10.2	数据来源与思路	139
10.3	描述性分析	139
10.4	时间序列趋势图	141
10.5	对数据进行相关分析	144
10.6	时间序列的单位根检验	145
10.7	两时间序列分析的协整检验与误差修正模型	151
10.8	格兰杰因果关系检验	154



练习题 .....	156
<b>第 11 章 时间序列分析 GARCH 模型的 R 语言应用 .....</b>	<b>157</b>
11.1 GARCH 模型的含义 .....	157
11.2 ARCH 效应检验 .....	157
11.3 GARCH 模型的 R 语言函数用法 .....	158
11.4 GARCH 模型的 R 语言函数应用实例 .....	158
11.5 德国股票指数的 GARCH 模型的 R 语言应用 .....	160
练习题 .....	163
<b>第 12 章 面板数据分析的 R 语言应用 .....</b>	<b>164</b>
12.1 面板数据分析的基本理论 .....	164
12.2 面板数据格式定义 .....	165
12.3 混合估计回归模型 R 语言估计 .....	166
12.4 固定效应回归模型 R 语言估计 .....	167
12.5 固定效应回归模型与混合估计回归模型优劣判断的 R 语言应用 .....	168
12.6 随机效应回归模型的 R 语言估计 .....	168
12.7 组间计量回归分析的 R 语言估计 .....	169
12.8 随机效应回归模型与固定效应回归模型区分的 Hausman 检验 .....	169
12.9 面板数据的广义矩估计的 R 语言应用 .....	170
12.9.1 资本资产定价模型检验的广义矩估计法(GMM)的 R 语言应用 .....	170
12.9.2 多因素套利定价模型检验的广义矩估计的 R 语言应用 .....	171
练习题 .....	173
<b>第 13 章 基于 R 语言的金融数据分析综合应用 .....</b>	<b>174</b>
13.1 构建金融数据分析平台的 R 程序包功能及层次 .....	174
13.2 数据处理和图形展示程序包 quantmod .....	174
13.3 金融数据获取 .....	179
13.4 时间序列分析的 R 工具 .....	181
13.5 时间序列分析 .....	183
13.6 金融数据分析 R 语言综合应用 .....	184
练习题 .....	189
<b>第 14 章 创业板科技型上市公司股权激励对其价值影响的计量检验研究 .....</b>	<b>190</b>
14.1 科技型上市公司股权激励相关概念 .....	190
14.1.1 科技型上市公司的概念与特征 .....	190
14.1.2 股权激励的概念及特点 .....	191
14.2 科技型上市公司股权激励的定性分析 .....	193

14.2.1	科技型上市公司推行股权激励的意义 .....	193
14.2.2	高新技术企业实施股权激励的可行性 .....	193
14.2.3	我国科技型上市公司股权激励现状分析 .....	195
14.2.4	股权激励实施效果分析 .....	199
14.3	科技型上市公司股权激励绩效的实证研究 .....	199
14.3.1	研究假设 .....	199
14.3.2	变量选取 .....	200
14.3.3	样本的选择和数据的来源 .....	200
14.3.4	模型设计 .....	201
14.3.5	实证检验 .....	201
14.3.6	小结 .....	206
14.4	研究结论及建议 .....	208
14.4.1	研究结论 .....	208
14.4.2	政策建议 .....	209
	练习题 .....	211
	参考文献 .....	212



# 经济与金融计量学绪论

本章简要介绍经济与金融计量的方法和一般应用步骤,着重介绍经济与金融数据的类型和特点,简要评述主要的计量和统计软件包,本章旨在使学生理解经济与金融计量模型思想,了解经济与金融数据的特点与来源,掌握常用的金融计量软件 R 语言。

## 1.1 经济计量学与金融计量学的含义及建模步骤

### 1.1.1 计量经济学与金融计量经济学的含义

计量经济学是将经济理论实用化、数量化的实证经济学,可简称为“经济中的测量”。它是利用经济理论、数学、统计推断等工具对经济现象进行分析的经济学科的分枝,具体包括模型设计和建立、参数估计和检验以及利用模型进行预测等过程。

自 1926 年挪威经济学家费里希首次提出计量经济学的概念以来(他仿照生物计量学一词提出计量经济学概念,并将其定义为统计学、经济学和数学的结合),计量经济学的建立到现在不到 100 年,但是这门学科已经得到广泛发展。截至 2008 年所产生的 61 位诺贝尔经济学奖获得者中,30 多位在获奖成果中应用了计量经济学。尤其是 20 世纪 90 年代以来,赫克曼、麦克法登、格兰杰、恩格尔等学者都是因为对计量经济学方面的突出贡献而获得诺贝尔奖的。

对于金融计量学的含义,在西方一般是指金融市场的计量分析,主要包括对金融市场各种变量(利率、汇率、交易量、价格等)进行相应的统计分析和计量建模,以及对实证金融中的大量金融理论和现象进行分析。

本书包括经济计量和金融计量两大部分。具体包括经济与金融计量学中经典回归模型及其应用、非经典回归模型及其应用(多重共线性、自相关、异方差)、时间序列分析及其应用、条件异方差模型及其应用、面板数据分析模型及其应用等,经济计量学部分主要介绍经典回归模型及其应用、非经典回归模型及其应用(多重共线性、自相关、异方差)。金融计量学部分主要介绍时间序列分析及其应用、条件异方差模型及其应用、面板数据分析模型及其应用等。

### 1.1.2 经济计量学与金融计量学建模过程

基于对经济计量学与金融计量学范畴的理解,对经济计量学与金融计量学建模的步骤描述如下:

步骤 1,关于研究问题的概述。该步骤通常涉及经济与金融理论的形成,或者来自某种理论的认识——两个或多个变量之间特定方式的联系。这一步需要将经济与金融理论或相

关变量之间的关系模型用数学的方式表达出来。

步骤 2, 样本数据收集。这一步骤是经济与金融计量工作的基础工作, 也是直接影响到检验结果的一项工作。通常应根据研究对象, 进行样本数据的收集和整理, 并在此基础上取舍变量, 并分析经济与金融数据的类型、特点和来源等。

步骤 3, 选择合适的计量方法来估计模型。在经济与金融计量过程中, 有必要根据研究目的以及数据本身的特点、需要, 选择相应的估计方法和计量模型, 如根据数据是连续数据还是离散数据选择一元回归、多元回归模型还是离散模型, 根据数据是一元时间序列还是多元时间序列选择相应的计算模型, 研究金融市场的波动率、利率期限结构则要选择相应的计量模型等。

步骤 4, 对模型进行实证检验。在估计参数后, 一个初步的模型就构建起来了, 但是为了评估所建立的模型是否合适, 能否反映变量之间的关系, 还需要对模型进行进一步检验。模型检验通常包含统计检验、计量经济学检验以及经济金融意义检验三方面的内容。统计检验的目的在于检验模型参数估计值的可靠性, 包括模型的拟合度检验、变量的显著性检验等; 计量经济学检验是根据计量经济学理论的要求而进行的, 包括序列相关性检验、异方差性检验和多重共线性检验等; 经济金融意义检验是将计量检验的结果与相应的经济理论或金融理论比较以检验是否相符。若所构建的模型估计结果不能通过上述某方面的检验, 就有必要考虑前面几个步骤中是否存在问题并重新建立模型; 若能够通过模型的检验, 则可进入经济与金融计量模型的应用阶段。

步骤 5, 模型应用。若模型能够通过检验, 则说明所构建的计量模型具有适用性, 这样就可以将模型应用于特定的研究。通常, 所构建的模型主要有以下三个方面的应用:

- (1) 结构分析, 即研究一个变量或几个变量变化时对其他变量的影响, 以揭示不同经济变量之间的内在联系;
- (2) 金融经济预测, 即根据金融经济模型对未来金融经济变量的变化进行预测分析;
- (3) 政策评价, 即研究不同的政策对经济目标所产生影响的差异, 从金融计量分析中寻求优化政策目标的路径。

经济与金融计量建模的基本步骤如图 1-1 所示。

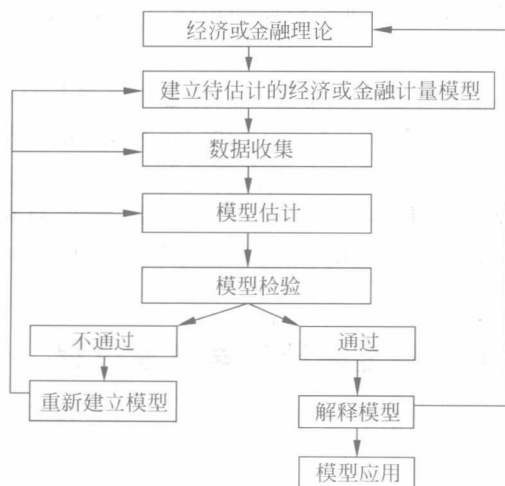


图 1-1 经济与金融计量建模的基本步骤

### 1.1.3 经济与金融模型中的数据

从构建经济与金融计量模型的步骤来看,数据分析是重要的环节。下面,着重分析经济与金融数据的类型、特点和来源。

#### 1. 经济与金融数据的类型

经济与金融计中需要处理的数据主要有三类:横截面数据、时间序列数据和面板数据。

(1) 横截面数据。是同一时间(时期或时点)某一指标在不同空间的观测数据。如某一时点中国 A 股市场的平均收益率,2010 年所有 A 股上市公司的净资产收益率。在利用横截面数据进行分析时,由于单个或多个解释变量观测值起伏变化会对被解释变量产生不同的影响,因而导致异方差问题。因此在数据整理时必须消除异方差。

(2) 时间序列数据。即按时间序列排列的数据,也称为动态序列数据。时间序列数据是按照一定时间间隔对某一变量或不同时间的取值进行观测所得到的一组数据,例如每一季度的 GDP 数据、每一天的股票交易数据或债券收益率数据等。在金融计量分析中,时间序列数据是常见的一类数据类型。

(3) 面板数据。即时间序列数据和横截面数据相结合的数据。

金融领域以时间序列数据与面板数据为主。

#### 2. 经济与金融数据来源

经济与金融数据主要有以下三个来源:

(1) 专业性网站。如国家统计局网站、中国人民银行网站、中国证监会网站、世界银行网站、国际货币基金组织网站等。

(2) 专业数据公司和信息公司。国外数据库主要有芝加哥大学商学院的证券价格研究中心(CRSP)、路透(Reuter)终端等。国内的经济与金融数据库主要有 CCER 中国金融经济数据库、国泰安数据库(GTA)、万德数据库(Wind)、锐思数据库等。如表 1-1 所示。

表 1-1 金融数据库

数据库名称	网 址
CRSP	www. chicagobooth. edu
Reuter	www. Reuters. com
Bloomberg	www. Bloomberg. com
Wind	www. wind. com. cn
GTA	www. gtadata. com
CCER 中国金融经济数据库	www. ccer. edu. cn
聚源锐思数据库	www. resset. cn
天相金融数据库	www. txsec. com/zqsc/tx_data. asp
万得 Wind 金融数据库	www. wind. com. cn

(3) 抽样调查。是针对某些专门的研究开展的一类获取数据的方式。比如,要对中国的投资者信心进行建模,就必须通过设计调查问卷,对不同的投资群体进行数据采集。

## 1.2 经济与金融计量软件简介

### 1.2.1 R 软件简介

R 语言是统计领域广泛使用的诞生于 1980 年左右的 S 语言的一个分支,可以认为 R 语言是 S 语言的一种实现。而 S 语言是由 AT&T 贝尔实验室开发的一种用来进行数据探索、统计分析和作图的解释型语言。最初 S 语言的实现版本主要是 S-PLUS。S-PLUS 是一个商业软件,它基于 S 语言,并由 MathSoft 公司的统计科学部进一步完善。后来 Auckland 大学的 Robert Gentleman 和 Ross Ihaka 及其他志愿人员开发了一个 R 系统。R 语言是基于 S 语言的一个 GNU 项目,所以也可以当作 S 语言的一种实现,通常用 S 语言编写的代码都可以不作修改地在 R 语言环境下运行。R 语言的语法来自 Scheme。R 的使用与 S-PLUS 有很多类似之处,这两种语言有一定的兼容性。S-PLUS 的使用手册只要稍加修改就可作为 R 语言的使用手册。所以有人说:R 是 S-PLUS 的一个“克隆”。

详细内容请登录 <http://cran.r-project.org> 查询。

### 1.2.2 Python 软件简介

Python 是一种面向对象的解释型计算机程序设计语言,由 Guido van Rossum 于 1989 年底发明,第一个公开发行人版发行于 1991 年。Python 源代码同样遵循 GPL(GNU General Public License)协议。Python 语法简洁而清晰,具有丰富和强大的类库。它常被昵称为胶水语言,能够把用其他语言制作的各种模块(尤其是 C/C++)很轻松地联结在一起。常见的一种应用情形是,使用 Python 快速生成程序的原型(有时甚至是程序的最终界面),然后对其中有特别要求的部分用更合适的语言改写,比如,3D 游戏中的图形渲染模块性能要求特别高,就可以用 C/C++ 重写,而后封装为 Python 可以调用的扩展类库。需要注意的是,在使用扩展类库时可能需要考虑平台问题,某些扩展类库可能不提供跨平台的实现。

Python 需要安装 numpy、pandas、scipy、cython、statsmodels、matplotlib 等一系列的程序包,还需要安装 ipython 交互环境,单独用 Python 直接做计量分析统计函数是没有函数支持的。

详细内容请登录 <https://www.python.org/> 查询。

### 1.2.3 Stata 软件简介

Stata 由美国计算机资源中心(Computer Resource Center)于 1985 年研制。其特点是采用命令行/程序操作方式,程序短小精悍,功能强大。Stata 是一套提供了数据分析、数据管理以及绘制专业图表功能的完整及整合性统计软件。它提供的功能包含线性混合模型、均衡重复反复及多项式普罗比模式。新版本的 Stata 采用最具亲和力的图形界面,使用者自行建立程序时,软件能提供具有直接命令式的语法。Stata 提供完整的使用手册,包含统计样本建立、解释、模型与语法、文献等。

Stata 目前最新版为 13 版。

详细内容请登录 <http://www.stata.com> 查询。

### 1.2.4 EViews 软件简介

EViews 是美国 GMS 公司于 1981 年发行的 Micro TSP 的 Windows 版本,通常称为计量经济学软件包。EViews 是 Econometrics Views 的缩写,它的本意是对社会经济关系与经济活动的数量规律,采用计量经济学方法与技术进行“观察”。计量经济学的核心是设计模型、收集资料、估计模型、检验模型、运用模型进行预测、求解模型和应用模型。EViews 是完成上述任务必不可少的工具。正是由于 EViews 等计量经济学软件包的出现,使计量经济学取得了长足的进步,发展成为实用与严谨的经济学科。使用 EViews 软件包可以对时间序列和非时间序列的数据进行分析,建立序列(变量)间的统计关系式,并用该关系式进行预测、模拟等。虽然 EViews 是由经济学家开发的,并且大多应用于经济学领域,但并不意味着必须限制该软件包仅只用于处理经济方面的时间序列。EViews 处理非时间序列数据照样得心应手。实际上,大型的非时间序列(截面数据)的项目也能在 EViews 中进行处理。

详细内容请登录 <http://www.eviews.com/> 查询。

### 1.2.5 SAS 软件简介

SAS 是美国 SAS 软件研究所研制的一套大型集成应用软件系统,具有完备的数据存取、数据管理、数据分析和数据展现功能。尤其是创业产品统计分析系统部分,由于其具有强大的数据分析能力,一直是业界著名软件,在数据处理和统计分析领域被誉为国际上的标准软件和最权威的优秀统计软件包,广泛应用于政府行政管理、科研、教育、生产和金融等领域,发挥着重要的作用。SAS 系统中提供的主要分析功能包括统计分析、经济计量分析、时间序列分析、决策分析、财务分析和全面质量管理工具等。

详细内容请登录 <http://www.sas.com> 查询。

### 1.2.6 Matlab 软件简介

Matlab 软件是由美国 Mathworks 公司推出的用于数值计算和图形处理的科学计算系统,在 Matlab 环境下,用户可以集成地进行程序设计、数值计算、图形绘制、输入输出、文件管理等各项操作。它提供的是一个人机交互的数学系统环境,与利用 C 语言作数值计算的程序设计相比,利用 Matlab 可以节省大量的编程时间,且程序设计自由度大。其最大的特点是直观、简洁的程序开发环境,语言简洁紧凑,使用方便灵活,库函数与运算符极其丰富,另外它具有强大的图形功能。

在国际学术界,Matlab 已经被确认为准确、可靠的科学计算标准软件,在许多国际一流学术刊物上都可以看到 Matlab 的应用。

详细内容请登录 <http://www.mathworks.com> 查询。

### 1.2.7 SPSS 软件简介

SPSS(Statistical Package for the Social Science,社会科学统计软件包)是世界著名的统计分析软件之一。20 世纪 60 年代末,美国斯坦福大学的三位研究生研制开发了最早的统计分析软件 SPSS,同时成立了 SPSS 公司,并于 1975 年在芝加哥组建了 SPSS 总部。20

世纪 80 年代以前,SPSS 统计软件主要应用于企事业单位。1984 年 SPSS 总部首先推出了世界第一个统计分析软件微机版本 SPSS/PC+,开创了 SPSS 微机系列产品的开发方向,从而确立了个人用户市场第一的地位。2009 年 IBM 公司收购 SPSS 公司后,在中国国内市场推出的最新产品是 IBM SPSS Statistics 21.0 多国语言版。SPSS/PC+ 的推出极大地扩充了它的应用范围,使其能很快地应用于自然科学、技术科学、社会科学的各个领域。它使用 Windows 的窗口方式展示各种管理和分析数据的功能,使用对话框展示各种功能选择项,只要掌握一定的 Windows 操作技能,粗通统计分析原理,就可以使用该软件为特定的科研工作服务。

详细内容请登录 <http://www.spss.com> 查询。

还有一些统计和计量经济学软件,如 Statistica、S-PLUS 等,但相对来说没有上面 6 种软件流行。各软件网站列表如表 1-2 所示。

表 1-2 金融经济计量软件网站

软件名称	网 址
R	<a href="http://www.cran.r-project.org">www.cran.r-project.org</a>
Stata	<a href="http://www.stata.com">www.stata.com</a>
EViews	<a href="http://www.eviews.com">www.eviews.com</a>
SAS	<a href="http://www.sas.com">www.sas.com</a>
Matlab	<a href="http://www.mathworks.com">www.mathworks.com</a>
SPSS	<a href="http://www.spss.com">www.spss.com</a>
S-PLUS	<a href="http://www.mathsoft.com">www.mathsoft.com</a>
Statistica	<a href="http://www.statsoft.com">www.statsoft.com</a>

## 练习题

1. 什么是经济计量学和金融计量学?
2. 简述经济和金融计量的常用软件。





# R 语言的下载、安装与启动

## 第 2 章

### 2.1 选择 R 语言的理由

R 语言是一个有着统计数据分析功能及强大作图功能的软件系统,是一种新兴的统计学软件、语言和环境,而且它的源代码开放。R 语言以其强大功能和它在统计理论及分析上的优势,近年来在统计、经济、管理、金融等相关领域受到有关人士的广泛欢迎和关注。虽然我使用过 Matlab、SAS、SPSS 和 Stata 等统计计算方面的软件,但现在 R 语言是我的首选。原因如下:

- R 语言是自由免费软件。它不收取任何费用,但其能力不会比任何同类型的商业软件差。从功能相似的角度来说,R 和 Matlab 是最像的。
- 通过 R 语言,可以和全球一流的统计计算方面的专家进行讨论,它是全世界统计学家思维的最大集中。
- R 语言是彻底的面向对象的统计编程语言。对于熟悉面向对象编程的人来说非常容易理解和使用。
- R 语言和其他编程语言/数据库之间有很好的接口。代码整合的时候感觉 R 语言为用户提供了一系列对象,用其他语言只要调用这些对象就可以了,这对数据整合工作非常有用。
- R 语言浮点运算功能强大。R 语言可以作为一台高级科学计算器,因为 R 同 Matlab 一样不需要编译就可执行代码。
- R 语言不依赖于操作系统。R 语言可以在运行于 Windows、UNIX、Linux 和 Macintosh 等操作系统上,它们的安装文件以及安装说明都可以在 CRAN 社区上下载。
- R 语言的帮助功能完善。R 语言嵌入了非常实用的帮助系统,这个帮助系统随软件所附的 pdf 或 html 帮助文件可以随时通过主菜单打开浏览或打印,通过 help 命令可随时了解 R 语言所提供的各类函数的使用方法和例子。
- R 语言作图功能强大。R 语言内嵌的作图函数能将产生的图片展示在一个独立的窗口中,并能将之保存为各种形式的文件(如 jpg、png、bmp、ps、pdf、emf、pictex、xfig 等)。
- R 语言统计分析能力尤为突出。R 语言内嵌了许多实用的统计分析函数,统计分析的结果也能被直接显示出来,一些中间结果(如 p 值、回归系数、残差等)既可保存到专门的文件中,也可直接用于进一步分析。R 语言的部分统计功能整合在 R 语言的