



海南大学经济管理系列丛书

蒙特卡洛模拟及其Stata 应用实现

*MengTeKaLuo MoNi JiQi Stata
YingYong ShiXian*

陈国汉 著



经济科学出版社
Economic Science Press



海南大学经济管理系列丛书

蒙特卡洛模拟及其Stata 应用实现

*MengTeKaLuo MoNi JiQi Stata
YingYong ShiXian*

陈国汉 著



经济科学出版社
Economic Science Press

图书在版编目 (CIP) 数据

蒙特卡洛模拟及其 Stata 应用实现 / 陈国汉著. —北京:
经济科学出版社, 2015. 10

(海南大学经济管理系列丛书)

ISBN 978 - 7 - 5141 - 6101 - 4

I. ①蒙… II. ①陈… III. ①蒙特卡洛法 - 应用 -
测量 IV. ①TB9

中国版本图书馆 CIP 数据核字 (2015) 第 231476 号

责任编辑: 刘明晖 李 军

责任校对: 王肖楠

版式设计: 齐 杰

责任印制: 王世伟

蒙特卡洛模拟及其 Stata 应用实现

陈国汉 著

经济科学出版社出版、发行 新华书店经销

社址: 北京市海淀区阜成路甲 28 号 邮编: 100142

总编部电话: 010 - 88191217 发行部电话: 010 - 88191540

网址: www.esp.com.cn

电子邮件: esp@esp.com.cn

天猫网店: 经济科学出版社旗舰店

网址: <http://jjkxcbs.tmall.com>

北京盛源印刷有限公司印装

710 × 1000 16 开 11 印张 170000 字

2015 年 11 月第 1 版 2015 年 11 月第 1 次印刷

ISBN 978 - 7 - 5141 - 6101 - 4 定价: 32.00 元

(图书出现印装问题, 本社负责调换。电话: 010 - 88191502)

(版权所有 翻印必究 举报电话: 010 - 88191586)

电子邮箱: dbts@esp.com.cn

前 言

蒙特卡洛方法被人们称为“最后的方法”，因为它可以解决其他数值方法不能解决的问题。人们所面临的大量实际问题中都存在不确定性，一般通过建立数学模型（主要是统计模型）来概括实际问题，但是许多问题由于本身的原因，并不适合建立模型，或者建立的模型过于复杂无法求的解析解。在这种情况下蒙特卡洛模拟方法应运而生。蒙特卡洛方法正式的名称是随机模拟，由于其命名的传奇色彩，所以现在一般以蒙特卡洛法代指随机模拟。它通过简单的手段仿真随机系统的运行来获得系统的状态变化与输出结果的大量数据，以解决实际问题。它简单实用，且实用面广，是当今科学技术研究领域的有力的研究手段。

本书的第1章是 Stata 软件基础，主要介绍了 Stata 软件的一些基本功能与操作。第2章介绍了 Stata 软件的语法结构，尤其着重介绍了蒙特卡洛模拟的相关命令。第3章简单地介绍了 Stata 软件的数据管理功能。第4章介绍了蒙特卡洛方法的基础，包括其基本思想，基本特点，随机性与随机数，随机数发生器以及蒙特卡洛方法的发展简史。第5章介绍了伪随机数的生成，介绍了几种比较常用的伪随机数发生器以及伪随机数的检验方法。第6章介绍了概率分布的抽样方法，介绍了直接抽样原理，随机变量的几种常用抽样方法，随机向量抽样方

法，随机过程抽样方法以及未知概率分布抽样方法。第7章介绍了离散型随机变量的模拟，主要介绍了二项分布，泊松分布以及未知事件空间有限的离散型随机变量的模拟。第8章介绍了连续型随机变量的模拟，介绍了均匀分布、正态分布等几种主要的连续性随机变量分布的模拟，以及多维正态分布的模拟。第9章介绍了随机过程的模拟，主要介绍了平稳随机过程以及布朗运动的模拟。第10章是蒙特卡洛模拟的应用，介绍了不同领域中的一些蒙特卡洛模拟的实例。

本书把程序设计方法和蒙特卡洛模拟方法融合在一起，目的是使读者在学习蒙特卡洛方法的同时即可使用 Stata 编程来进行模拟练习，相信会给读者带来事半功倍的效果。本书适用于科研人员、研究生，以及高年级本科生，既可以作为教学辅助材料，也可用于蒙特卡洛模拟的教学参考。本书中难免有不足之处，敬请读者批评指正。

陈国议

2015年9月

目 录

第 1 章 Stata 软件基础	1
1.1 Stata 软件简介	1
1.2 Stata 窗口及基本操作	2
1.3 Stata 帮助与支持	5
第 2 章 Stata 的语法结构	7
2.1 命令的基本格式	7
2.2 命令的示例	9
2.3 蒙特卡洛模拟命令	14
2.4 绘图命令	17
第 3 章 Stata 的数据管理	19
3.1 变量的基本操作	19
3.2 数据的基本操作	26
3.3 日期与时间函数	32
第 4 章 蒙特卡洛方法基础	34
4.1 蒙特卡洛方法的基本思想	34
4.2 蒙特卡洛方法的基本特点	36
4.3 随机性与随机数	38
4.4 随机变量的抽样	43
4.5 蒙特卡洛方法简史	43

第 5 章	伪随机数的生成	45
5.1	伪随机数发生器简介	45
5.2	均匀随机数的产生	49
5.3	均匀随机数的检验	56
第 6 章	概率分布抽样方法	61
6.1	直接抽样方法原理	61
6.2	随机变量基本抽样方法	61
6.3	随机向量抽样方法	65
6.4	随机过程抽样方法	67
6.5	未知概率分布抽样方法	71
第 7 章	离散型随机变量的模拟	74
7.1	离散型随机变量	75
7.2	二项分布	79
7.3	几何分布	84
7.4	泊松分布	89
第 8 章	连续型随机变量的模拟	93
8.1	均匀分布	93
8.2	正态分布	98
8.3	多维正态分布	110
8.4	截断分布抽样	112
第 9 章	随机过程的模拟	114
9.1	平稳随机过程的模拟	115
9.2	布朗运动的模拟	123
第 10 章	蒙特卡洛模拟应用	131
10.1	简单概率问题求解	131

10.2	期权定价问题	142
10.3	随机服务问题	150
10.4	风险分析问题	157
10.5	积分计算问题	161
	参考文献	165
	后记	166

第 1 章

Stata 软件基础

Stata 与 SPSS、SAS 并称为当今三大统计软件。与后者相比，Stata 体积小、简单易懂且功能强大。Stata 把 EViews，SPSS 的傻瓜式菜单和 SAS 的命令、编程完美结合起来，所以它一推出就受到了初学者和高级用户的普遍欢迎。Stata 不仅在统计方面功能齐全，其在计量分析领域更是有着深刻影响，目前，Stata 是计量经济学，特别是微观计量经济学的主流软件。

1.1 Stata 软件简介

Stata 是一个用于数据管理和分析的统计软件。其最初由美国计算机资源中心（Computer Resource Center）研制，由 stata 公司（stata-Corp）开发。从 1985 年开始，先后开发了多个版本、日益完善目前 Stata14.0 已经发布。Stata 相比于 SAS 操作简便、易学易用，能够非常快捷轻便地进行统计和计量分析，与 SAS、SPSS 并称为三大权威统计软件。Stata 的数据格式简单，分析结果输出简捷易读，而且编程方便、绘图精美、运算速度快，适用于统计的教学。

Stata 主要用于数据处理和绘图、统计分析、调查分析与生存分析、抽样和计算机模拟、计量模型、编程等方面，以其强大的统计和计量功能广受用户推崇。其主要功能有：（1）单元统计：描述统计、假设检验（参数、非参数）、ANOVA、质量控制、统计作图等。（2）多元统

计：MANOVA、主成分、因子分析、典型相关、聚类、判别分析、对应分析、多维标度分析、线性回归、非线性回归、工具变量回归、广义线性回归、分位数回归（稳健回归）、系统方程模型（SUR、联立方程）、离散选择模型（二项选择、排序选择、多项选择、条件 Logit、嵌套 Logit 模型、二元选择模型等）、计数模型（泊松回归、负二项回归）、截断与归并模型、海克曼选择模型、逐步回归（stepwise）等。时间序列分析：时间序列的平滑、相关图、ARIMAX、GARCH、单位根检验、Johansen 协整检验、VAR、VEC、滚动回归等。（3）面板数据 [线性模型、工具变量回归、动态面板、分层混合效应、广义估计方程（GEE）、随机边界模型等]。（4）调查数据分析。（5）生存分析。（6）蒙特卡洛模拟、自举法（bootstrap）、刀切法（jackknife）等。

Stata 的高级统计模块大多是编程人员用宏语言编写成的程序语言（ado 文件），这些文件可由用户自行修改、添加和下载。用户可以到 Stata 网站搜索下载最新的升级文件。这个特点使得 Stata 始终走在统计分析方法的最前沿，用户能够快速找到最新统计算法的 Stata 程序版本，也使得 Stata 成为统计软件中升级最多、最频繁的一个。

1.2 Stata 窗口及基本操作

运行 Stata 后，将会看到如图 1-1 所示的操作界面。

从图 1-1 中可以看出，Stata 最主要的部分是由四大窗口组成的，它们是分别是命令回顾窗口（Review）、结果窗口（Results）、变量窗口（Variables）和命令输入窗口（Command），接下来，将会详细地介绍一下这四个窗口。

（1）命令回顾窗口（Review），主要是用来临时性存储已经执行过的命令语句的窗口，这些执行的命令语句主要包括两种：一种是直接从命令窗口中输入的命令，另一种是通过窗口菜单操作转化而成的命令。Review 窗口可以临时性存储自 Stata 本次运行到结束的所有命令，若 Stata 中途或最终被关闭，则所有的命令语句将会自动消失，若想永久保存这些命令，可以通过使用 log 命令或单击右键实现存储。在 Stata 运

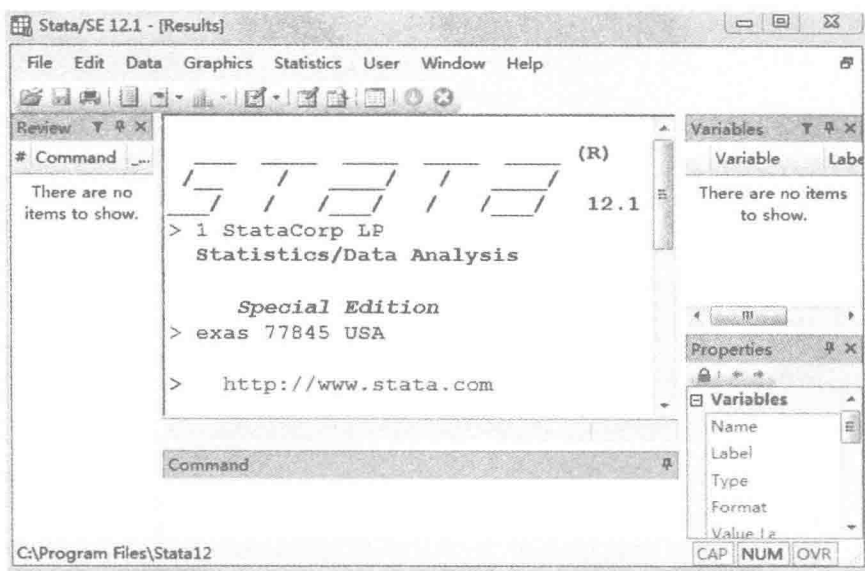


图 1-1 Stata12.1 启动后的界面

行过程中，可以重复使用显示在 Review 窗口中的命令，只需要左键单击命令，该命令将会重新显示在 Command 窗口中，供用户进行修改和执行。

(2) 结果窗口 (Results)，主要是用来显示命令执行结果的窗口。若是命令执行的结果过长，则会在命令窗口的底端出现“more”这一字样，这时只需要按下空格键或者回车键，就可以浏览下面的内容；如果想停止浏览，只需要按字母“q”键就可停止。

(3) 变量窗口 (Variables)，主要是用来显示变量名称和类型的窗口。点击某个变量名称，其就可以显示在右方的命令窗口中。如果想要删除某个变量，则只需要在命令窗口中输入 drop 命令，然后鼠标左键单击相应的变量名称，然后点击回车键即可完成删除操作。

(4) 命令输入窗口 (Command)，是用户进行交互式程序操作的主要实现场所，只要用户将相关命令输入，然后点击回车键即可。

工具栏是一些常用的快捷按钮，我们仅做一些必要的说明。

Save 按钮：用于保存当前正在使用的，存在于内存中的 Stata 数据，这些数据以变量的形式存在，该菜单命令将这些变量保存在一个 Stata

电子表格中。一般情况下不建议使用这一按钮，因为你的数据可能来自于一个已经存在的数据文件，单击 save 按钮以后，Stata 按照原来调用的地址和文件名进行保存，会覆盖原来的文件。在打开和保存之间，我们的操作可能已经修改过其中的数据，而这一过程是不可逆的，保存以后就找不到原始文件了。在实证研究中，往往需要我们保存最原始的数据和给出处理的过程，我们建议在命令窗口中输入 save 命令来保存。

Log 按钮：用来开启一个 log 文件。Log 文件可以保存在屏幕输出窗口中的结果，其具体操作我们将在下文介绍。

图形窗口前置：点击该按钮的前提条件是我们已经绘制了一个图形，或者通过命令调入了一个图形文件，否则该按钮将显示为灰色（非激活状态）。

数据编辑器：可以对数据进行录入、编码、修改等工作。

数据浏览器：在该窗口中只能进行数据查看，不能对数据进行编辑和修改。

程序终止：在执行程序的过程中，用户可能会忽然想到程序需要修改某些内容，或者由于编写上存在错误导致出现死循环，迟迟不能给出计算结果，这时需要点击该按钮。

在 Stata 中既可以使用类似 EViews、SPSS 的傻瓜式菜单的操作方式，也可以使用类似 SAS 的命令、编程的操作方式。

1. 菜单式操作。Stata 的菜单栏是用户进行菜单操作的主要媒介和工具。菜单栏主要包括 File、Edit、Data、Graphics、Statistics、User、Window、Help 这八个子菜单。

(1) File 的下拉菜单包括打开、保存、查看文件，导入、导出数据以及打印等功能。

(2) Edit 的下拉菜单包括数据的复制、粘贴等有关数据管理和设置的功能。

(3) Data 的下拉菜单包括数据的描述、编辑、浏览、增加或删除变量、文件合并、矩阵操作等方面的内容。

(4) Graphics 主要是用来作图的菜单，作图种类主要包括散点图、线图、柱状图、饼图等各种图形。

(5) Statistics 主要是用来进行各种统计和计量分析的菜单，主要包

括线性回归模型分析、时间序列分析、面板数据分析等方面的内容。

(6) User 主要是用来构建用户自己的菜单，主要包括有关数据、图表和统计分析等各方面的设置和操作。

(7) Window 主要是用于对显示界面的操作，主要包括对 Review、Results、Variables、Command 四大窗口的操作。

以多元线性回归为例，在菜单栏中选择 Statistics 选项，然后在里面选择相应的统计分析项目。此处做的是多元线性回归，那么在下拉菜单中选 Linear models and related 中的 Linear regression，然后填入相应的因变量 (price)、自变量 (mpg; weight) 即可。故只要知道统计项目的英文名称，那么使用起来还是很方便的。

2. 命令式操作。在 Stata 软件界面下方的 Command 命令窗口直接输入命令进行操作。由于 Stata 的命令简洁方便，所以很多用户在一开始就接受了这种操作方式。同时，在使用菜单式操作时，Stata 会在结果的上方输出对应的命令，所以这也帮助用户熟悉其命令格式。

3. Do-file 编程式操作，即将若干条命令组合在一起，按需求加上一定的循环、条件语句而组成的程序，并运行。编程式操作一般是用户对命令非常熟悉时，并要完成批量或更为复杂的任务时所采用的操作方式。但是由于程序有很强的移植性，所以可以从其他人的 Do-file 程序中获取自己想要的，并进行局部修改即可为自己所用，而不必自己重新写代码。同时，Do-file 文档可以保存为 *.do 的文本文件格式，方便修改和以后使用。由于此处只是一句命令语句，并不复杂，所以和命令式操作方式一致。

总结一下：这三种操作方式为初学者和高级用户均提供了合适的环境，一般是逐层递进的。但是我们首推后两种操作方式，因为它们效率高，并有很强的移植性，即简单修改方可为我们所用。而且 Do-file 编程式操作方便修改和保存。故我们后续介绍以后两种方式为主。

1.3 Stata 帮助与支持

Stata 为用户提供了强大的帮助系统，新用户可以通过帮助系统的

应用，更好地利用 Stata 完成自己所需要的功能和操作。Stata 的帮助系统主要由 Stata 手册、Stata 自带帮助和网络帮助三个方面组成。

(1) Stata 手册是一本学习 Stata 使用的权威书籍，它按字母顺序排列出了 Stata 所有相关的命令。

(2) Stata 自带帮助系统是使用最方便，也是最常用的方法。输入 help 命令显示出 Stata 所有帮助内容的目录结构。如果输入具体的命令，则只显示该命令的帮助，如

```
. help summarize
```

也可以通过菜单式的点选方式获得帮助：Help > > stata command... 在弹出的对话框中输入：summarize 然后回车，得到与 help summarize 同样的结果。

(3) Stata 的网络帮助系统更为强大，用户可以在网上查找 Stata 还没有内置化的命令，实现自行安装。Stata 的支持资源不仅来自 Stata 公司，也来自广大的 Stata 用户。Stata 的网址是 www.stata.com，它可以提供软件更新、常见问题解答等各项服务；Stata 的互联网论坛 Statistic Internet，专门的索引期刊 Stata Journal。网络帮助可以采用如下命令获得

```
. findit scat3,net
```

```
. search scat3,net
```

这两条命令等价，均为寻找绘三维立体图的命令 scat3。由于 scat3 不是 STATA 内置命令，所以需要通过这两个命令搜索并下载安装后才能使用。

help 和 search 都是查找帮助文件的命令，它们之间的区别在于 help 用于查找精确的命令名，而 search 是模糊查找。如果你知道某个命令的名字，并且想知道它的具体使用方法，只需在 Stata 的命令行窗口中输入 help 空格加上这个名字。回车后结果屏幕上就会显示出这个命令的帮助文件的全部内容。如果你想知道在 Stata 下做某个估计或某种计算，而不知道具体该如何实现，就需要用 search 命令了。使用的方法和 help 类似，只需把准确的命令名改成某个关键词。回车后结果窗口会给出所有和这个关键词相关的帮助文件名和链接列表。在列表中寻找最相关的内容，点击后在弹出的查看窗口中会给出相关的帮助文件。耐心寻找，反复实验，通常可以较快地找到你需要的内容。

第 2 章

Stata 的语法结构

Stata 的语法十分地简单和灵活。一个命令便构成了一条可执行的语句，在 Stata Command 中输入后，敲回车键就可以运行了。当然也可以将一组语句编成 do-file 文件，按前边介绍的方法运行就可以了。每个命令都有调用的格式，只要符合 Stata 规定的格式，语法上就是无误的。至于具体的格式，使用时查一下 Stata 的帮助系统或工具手册就可以了。本章将主要介绍一下 Stata 的语法结构。

2.1 命令的基本格式

Stata 命令可以大体上分成 4 类：处理和分析数据的命令；help 命令；针对 memory 的命令；数学命令。处理和分析数据的命令构成 Stata 命令的主体。它们可实现各种各样的功能，但大体具有如下的格式：

```
[ by varlist1 : ] command [ varlist2 ] [ = exp ] [ if exp ] [ in range ]  
[ weight ] [ using filename ] [ , options ]
```

说明，符号 [] 表示可选项。command 为 Stata 的命令函数，varlist 为变量，[if exp] 表示命令只针对满足 exp（一般是一个逻辑表达式）的观测；[in range] 表示命令只针对处在 range 指定的范围内的观测，如 in 5 指执行的范围是第 5 个观测，in -5 指执行的范围是倒数第 5 个观测，in 5/12 指执行的范围是从第 5 到第 12 个观测；[weight] 用于设定观测值的权重；[using filename] 表示使用的数据文件，[options] 表示命令的

选项,不同命令的选项也不同,根据情况和需要而定,Stata 的强大功能主要就体现在这些选项上,选项的使用非常精细和讲究,使用时既要依据个人的经验,也要参看 Stata 的帮助系统或工具手册; [by varlist1:] 表示按一组变量(由 varlist1 指定)分组,对每一组分别执行命令(command); command [varlist2] 表示命令是针对由 varlist2 指定的变量。

varlist 表示若干变量。对于数据中存在的变量,允许的表达形式包括 *、? 和 ~。其中,* 表示任意字符,? 表示一个字符,~ 表示两个变量之间的所有变量(根据数据中变量的存放位置)。比如,数据文件中共有 20 个变量,依次为 var1、var2、…、var20,则 var * 表示所有变量 var1-var20, var? 表示变量 var1、var2、…、var9, var1-var6 表示变量 var1、var2、…、var6。生成新变量时,变量名称不能简化。如果变量具有相同的前缀并且都以数字结尾,可以用 - 表示。比如,生成新变量 V1、V2、V3、V4,可以用如下两种形式。

```
. input v1 v2 v3 v4
```

```
. input v1-v4。
```

[if exp] 极大地体现了 Stata 的灵活性。逻辑表达式 exp 一般由以下成分构成:变量名;数字,字符,表示缺失值的“.”;关系运算符:==(等于),!=,~=(不等于),>(大于),<(小于),>=(大于或等于),<=(小于或等于);逻辑运算符:&(与),|(或),~(非)。以下是一些应用的例子:

```
if age >65 & age <85( age 大于 65,小于 85 的观测)
```

```
if place == "Canada" & pop ~=. (place 为"Canada"并且 pop 不缺失的观测)
```

```
if year == 1994 | year == 1997( year 为 1994 或 1997 的观测)
```

```
if ~( pop ==. & year ==. ) (排除 pop 和 year 都缺失的观测)
```

根据需要,逻辑表达式可以有非常复杂的形式。

[in range] 和 [if exp] 相当于从一个大样本中挑出符合条件的小样本,其用处体现在两个方面。首先是数据清理阶段,找出那些有明显的错误或有关缺失的观测。其次是在数据处理和分析阶段,找出有特别兴趣或意义的观测。

[weight] 用于设定观测值的权重,权数的类型包括:频率权数

(fweight 或 frequency)、抽样权数 (pweight)、解析权数 (aweight) 和重要性权数 (iweight)。频率权数即每个观测值的重复次数。抽样权数表示在抽样设计中观测值被抽取到的概率的倒数。解析权数与观测值的方差成反比, 比如第 j 个观测值的权数为 w_j , 则其方差为 σ^2/w_j 。重要性权数即表示观测值的重要性, 这种权数没有正式的统计定义。

[by varlist:] 表示对 varlist (分类变量) 中的每一类分别执行命令。by 是很多命令的前缀。事实上, Stata 的很多命令带有不同的前缀。比如, xi, simulate, statsby, rolling, bootstrap, jackknife, permute, svy, stepwise, capture, noisily, quietly, version 等。

自举法即通过有放回地重复抽样计算所需要的统计量, 并反复若干次以得到统计量的标准差或其他精度测度指标。命令前缀 bootstrap 命令格式为:

bootstrap exp_list[, options eform_option] : command

刀切法是 20 世纪 40 年代末 50 年代初发展起来的。它是指每一次删除样本中的一个观测值, 利用剩余的观测值进行估计, 进而得到所需要的统计量的分布。刀切法经常用于评估单个观测值对估计量的影响、计算统计量的标准差。命令前缀 jackknife 命令格式为:

jackknife exp_list[, options eform_option] : command

2.2 命令的示例

1. 命令 (command)

[by varlist:] **command** [varlist] [= exp] [if exp] [in range]
[weight] [, options]

查看 Stata 中所有命令的列表可以用如下命令:

```
. help command
```

打开 Stata 自身数据文件, 美国汽车数据 auto. dta:

```
. cd d: /stata12
```

```
. use auto, clear
```