



對外經濟貿易大學
UNIVERSITY OF INTERNATIONAL BUSINESS AND ECONOMICS

R语言

信息学院
汪波 副教授

2024年9月12日星期四

成绩构成

□ 40%平时成绩（出勤、作业）

- 考勤为不定期点名，大约每月一次（10%）
- 作业以数据分析题为主，题型与考试类似，可帮助大家熟悉期末考试（30%）

□ 60%期末考试

- 考试形式：开卷机考（统一在机房进行）
- 注意：考试以数据分析题为主，难度不大，但需要大家平时多熟悉操作命令，以免考试中临时查询花费时间

课程内容

- R语言的基本数据结构
 - 向量、数组与矩阵、列表与数据框
- R语言编程基础
 - 条件语句、循环语句、函数
- apply函数族
- 因子与分类汇总
- R语言画图
- 读写文件

什么是R?

The R Project for Statistical Computing

主页: <https://www.r-project.org/>

Getting Started

R is a **free** software environment for statistical computing and graphics. It compiles and runs on a wide variety of **UNIX platforms**, **Windows**, and **MacOS**. To **download R**, please choose your preferred **CRAN mirror**.

推荐参考书:

Robert I. Kabacoff, R语言实战（第二版）. 人民邮电出版社, 2016年5月

简史

R语言是从S统计绘图语言演变而来，可看作S的“方言”。

S语言上世纪70年代诞生于**贝尔实验室**，由Rick Becker, John Chambers, Allan Wilks开发。

基于S语言开发的商业软件**Splus**，可以方便的编写函数、建立模型，具有良好的扩展性，在国外学术界应用很广。

1995年由新西兰Auckland大学统计系的**R**obert Gentleman和**R**oss Ihaka，基于S语言的源代码，编写了能执行S语言的软件，并将该软件的源代码全部公开（**开源**），这就是R软件，其命令统称为R语言。

R与统计简介

很多人都把**R**作为一个统计系统来使用。我们倾向于把它当作**环境**，在功能上，**R**涵盖了绝大多数经典和现代统计技术。

一部分语句和程序已经被**内建**在基本的**R**语言环境中(**base**)，但是更多的是以**包(packages)**的形式提供的。

有**8**个包是随着**R**一同提供的（称作**标准包**），无需自行安装，但其它绝大多数的包要通过**CRAN**的成员网站获得并安装 (<http://cran.r-project.org>)。

R与其他主流的统计系统在本质上有个很重要的不同：在**R**中，统计分析通常由一系列步骤完成，同时将交互结果存储在**对象**中。因此，不同于**SAS**和**SPSS**在回归或者判别分析中会给出丰富的输出结果，**R**给出一个最小的输出，而将更多的结果保存在一个适当的**对象**中，由**R**函数进行后续的查询。

丰富的R语言资源

R源程序

不断更新：1997年 959K → 2013年 24M
→ 2020年(v4.0.0) 32M → 2023年 (v4.3.1) 79M

R核心团队

牛津大学、AT&T实验室等

R刊物

The R Journal

R书籍

上百种，包括统计编程，统计模型，数据分析，统计图示以及for Bioinformatics …

R会议

国际 - useR!; 中国 - COS (Chat Open Share)

R社区

R主页User Groups, 中文社区-统计之都

R的特点

多领域的统计资源

目前在R网站上有19828个程序包，涵盖了基础统计学、社会学、经济学、生态学、空间分析、系统发育分析、生物信息学、生物医学等诸多方面。

跨平台

R可在多种操作系统下运行，如Windows、MacOS、多种Linux和UNIX等。

命令行驱动

R是即时解释语言，输入命令，即可获得相应的执行结果。

为什么选择R?

□ 丰富的资源

涵盖了多种行业数据分析中几乎所有的办法。

□ 良好的扩展性

十分方便的编写函数和程序包，跨平台，可以胜任复杂的数据分析、绘制精美的图形。

□ 完备的帮助系统

每个函数都有统一格式的帮助，运行实例。

□ GNU软件

免费，软件本身及程序包的源代码均可公开获得。

R与其他统计软件比较

□ SAS

速度快，有大量统计分析模块，可扩展性稍差，昂贵。

□ SPSS

复杂的用户图形界面，简单易学，但不适合编程。

□ Splus

运行S语言，具有复杂的界面，与R完全兼容，昂贵。

R的缺点

□ 用户需要对命令熟悉

与代码打交道，需要记住常用命令。

□ 占用内存

所有的数据处理在内存中进行，不适于处理超大规模的数据。

□ 运行速度稍慢

即时编译，约相当于C语言的1/20。

□ 但相比点击鼠标的操作，R语言仍能大大提高效率。

CRAN

The Comprehensive R Archive Network

简称**CRAN**，由世界几十个镜像网站组成网络，提供下载安装程序和相应软件包。各镜像更新频率一般为1-2天
推荐镜像：

使用国内的镜像，如清华大学、中国科学技术大学等

<https://cran.r-project.org/mirrors.html>

Windows下载和安装R

- ❑ 要确保电脑的系统(如Windows)用户名是**英文**，否则R绘图会出现**乱码**而无法使用，后续安装RStudio也可能报错
- ❑ 推荐将R安装到软件的**默认目录**，否则后续安装RStudio时会无法直接找到R源文件
- ❑ CRAN: Binaries → Windows → base

Windows下载和安装R



[\[Home\]](#)

Download

[CRAN](#)

R Project

[About R](#)

[Logo](#)

[Contributors](#)

[What's New?](#)

[Reporting Bugs](#)

[Conferences](#)

[Search](#)

[Get Involved: Mailing Lists](#)

[Get Involved: Contributing](#)

[Developer Pages](#)

[R Blog](#)

R Foundation

[Foundation](#)

The R Project for Statistical Computing

Getting Started

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To [download R](#), please choose your preferred [CRAN mirror](#).

If you have questions about R like how to download and install the software, or what the license terms are, please read our [answers to frequently asked questions](#) before you send an email.

News

- useR! 2024 will be a hybrid conference, taking place 8-11 July 2024 in Salzburg, Austria.
- [R version 4.3.1 \(Beagle Scouts\)](#) has been released on 2023-06-16.
- [R version 4.2.3 \(Shortstop Beagle\)](#) has been released on 2023-03-15.
- You can support the R Foundation with a renewable subscription as a [supporting member](#)

News via Mastodon

<div class="loading-spinner"></div>

Windows下载和安装R

CRAN Mirrors

The Comprehensive R Archive Network is available at the following URLs, please choose a location close to you. Some statistics on the status of the mirrors can be found here: [main page](#), [windows release](#), [windows old release](#).

If you want to host a new mirror at your institution, please have a look at the [CRAN Mirror HOWTO](#).

0-Cloud

<https://cloud.r-project.org/>

Automatic redirection to servers worldwide, currently sponsored by Rstudio

Argentina

<http://mirror.fcaglp.unlp.edu.ar/CRAN/>

Universidad Nacional de La Plata

Australia

<https://cran.csiro.au/>

CSIRO

<https://mirror.aarnet.edu.au/pub/CRAN/>

AARNET

<https://cran.ms.unimelb.edu.au/>

School of Mathematics and Statistics, University of Melbourne

<https://cran.curtin.edu.au/>

Curtin University

Austria

<https://cran.wu.ac.at/>

Wirtschaftsuniversität Wien

Belgium

<https://www.freeststatistics.org/cran/>

Patrick Wessa

<https://ftp.belnet.be/mirror/CRAN/>

Belnet, the Belgian research and education network

Brazil

<https://cran-r.c3sl.ufpr.br/>

Universidade Federal do Parana

<https://cran.fiocruz.br/>

Oswaldo Cruz Foundation, Rio de Janeiro

<https://vps.fmvz.usp.br/CRAN/>

University of Sao Paulo, Sao Paulo

Windows下载和安装R

Canada

<https://mirror.rcg.sfu.ca/mirror/CRAN/>
<https://muug.ca/mirror/cran/>
<https://mirror.csclub.uwaterloo.ca/CRAN/>

Simon Fraser University, Burnaby
Manitoba Unix User Group
University of Waterloo

Chile

<https://cran.dcc.uchile.cl/>

Departamento de Ciencias de la Computación, Universidad de Chile

China

<https://mirrors.tuna.tsinghua.edu.cn/CRAN/>
<https://mirrors.bfsu.edu.cn/CRAN/>
<https://mirrors.pku.edu.cn/CRAN/>
<https://mirrors.ustc.edu.cn/CRAN/>
<https://mirrors.zju.edu.cn/CRAN/>
<https://mirror-hk.koddos.net/CRAN/>
<https://mirrors.e-ducation.cn/CRAN/>
<https://mirrors.qlu.edu.cn/CRAN/>
<https://mirror.lzu.edu.cn/CRAN/>
<https://mirrors.nju.edu.cn/CRAN/>
<https://mirrors.sjtug.sjtu.edu.cn/cran/>
<https://mirrors.sustech.edu.cn/CRAN/>
<https://mirrors.nwafu.edu.cn/cran/>

TUNA Team, Tsinghua University
Beijing Foreign Studies University
Peking University
University of Science and Technology of China
Zhejiang University
KoDDoS in Hong Kong
Elite Education
Qilu University of Technology
Lanzhou University Open Source Society
eScience Center, Nanjing University
Shanghai Jiao Tong University
Southern University of Science and Technology (SUSTech)
Northwest A&F University (NWAFU)

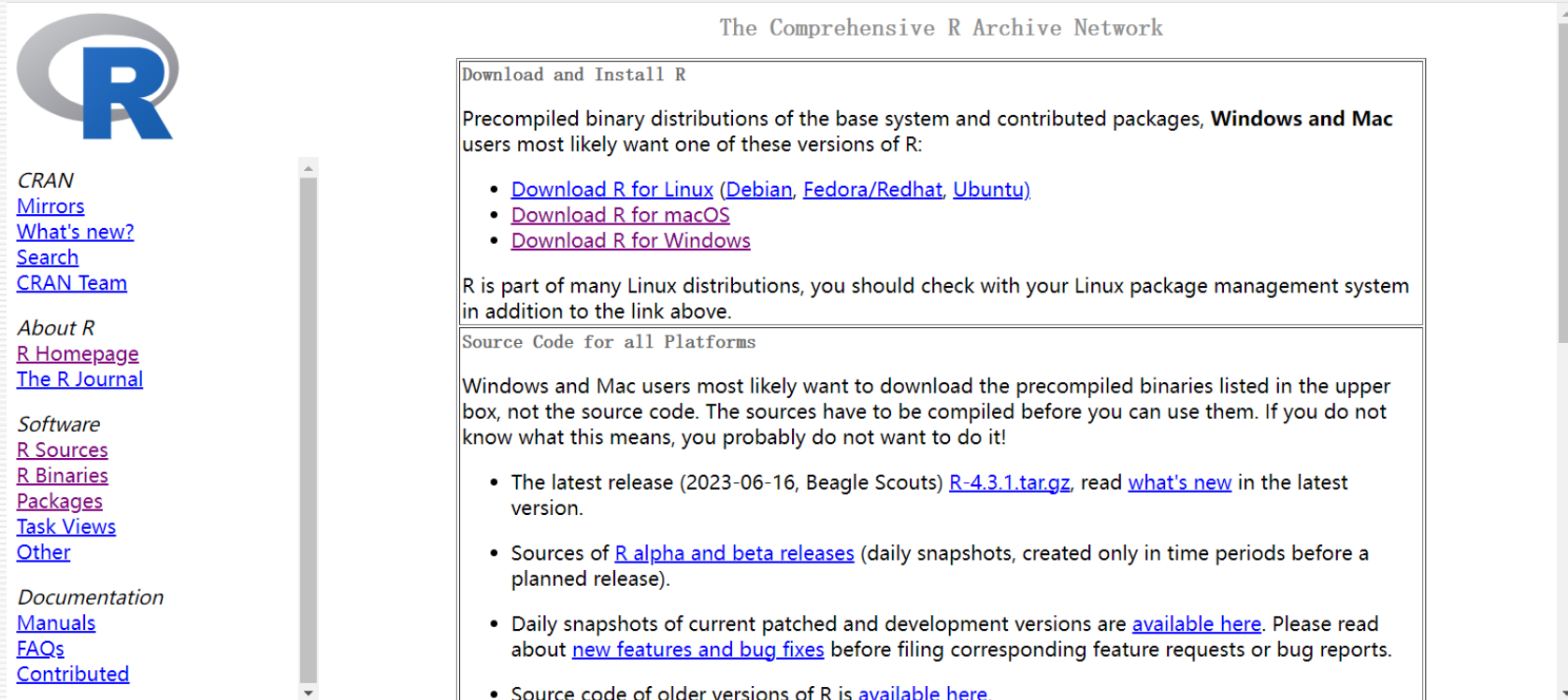
Colombia

<https://www.icesi.edu.co/CRAN/>

Icesi University

Costa Rica

Windows下载和安装R



The Comprehensive R Archive Network

Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- [Download R for Linux](#) ([Debian](#), [Fedora/Redhat](#), [Ubuntu](#))
- [Download R for macOS](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

Source Code for all Platforms

Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

- The latest release (2023-06-16, Beagle Scouts) [R-4.3.1.tar.gz](#), read [what's new](#) in the latest version.
- Sources of [R alpha and beta releases](#) (daily snapshots, created only in time periods before a planned release).
- Daily snapshots of current patched and development versions are [available here](#). Please read about [new features and bug fixes](#) before filing corresponding feature requests or bug reports.
- Source code of older versions of R is [available here](#).

CRAN
[Mirrors](#)
[What's new?](#)
[Search](#)
[CRAN Team](#)

About R
[R Homepage](#)
[The R Journal](#)

Software
[R Sources](#)
[R Binaries](#)
[Packages](#)
[Task Views](#)
[Other](#)

Documentation
[Manuals](#)
[FAQs](#)
[Contributed](#)

Windows下载和安装R



CRAN
[Mirrors](#)
[What's new?](#)
[Task Views](#)
[Search](#)

About R
[R Homepage](#)
[The R Journal](#)

Software
[R Sources](#)
[R Binaries](#)
[Packages](#)
[Other](#)

Documentation
[Manuals](#)
[FAQs](#)
[Contributed](#)

R for Windows

Subdirectories:

[base](#)

Binaries for base distribution. This is what you want to **install R for the first time**.

[contrib](#)

Binaries of contributed CRAN packages (for R \geq 2.13.x; managed by Uwe Ligges). There is also information on [third party software](#) available for CRAN Windows services and corresponding environment and make variables.

[old contrib](#)

Binaries of contributed CRAN packages for outdated versions of R (for R $<$ 2.13.x; managed by Uwe Ligges).

[Rtools](#)

Tools to build R and R packages. This is what you want to build your own packages on Windows, or to build R itself.

Please do not submit binaries to CRAN. Package developers might want to contact Uwe Ligges directly in case of questions / suggestions related to Windows binaries.

You may also want to read the [R FAQ](#) and [R for Windows FAQ](#).

Note: CRAN does some checks on these binaries for viruses, but cannot give guarantees. Use the normal precautions with downloaded executables.



對外經濟貿易大學
UNIVERSITY OF INTERNATIONAL BUSINESS AND ECONOMICS

Windows下载和安装R

R-4.3.1 for Windows

[Download R-4.3.1 for Windows](#) (79 megabytes, 64 bit)

[README on the Windows binary distribution](#)

[New features in this version](#)

This build requires UCRT, which is part of Windows since Windows 10 and Windows Server 2016. On older systems, UCRT has to be installed manually from [here](#).

If you want to double-check that the package you have downloaded matches the package distributed by CRAN, you can compare the [md5sum](#) of the .exe to the [fingerprint](#) on the master server.

Frequently asked questions

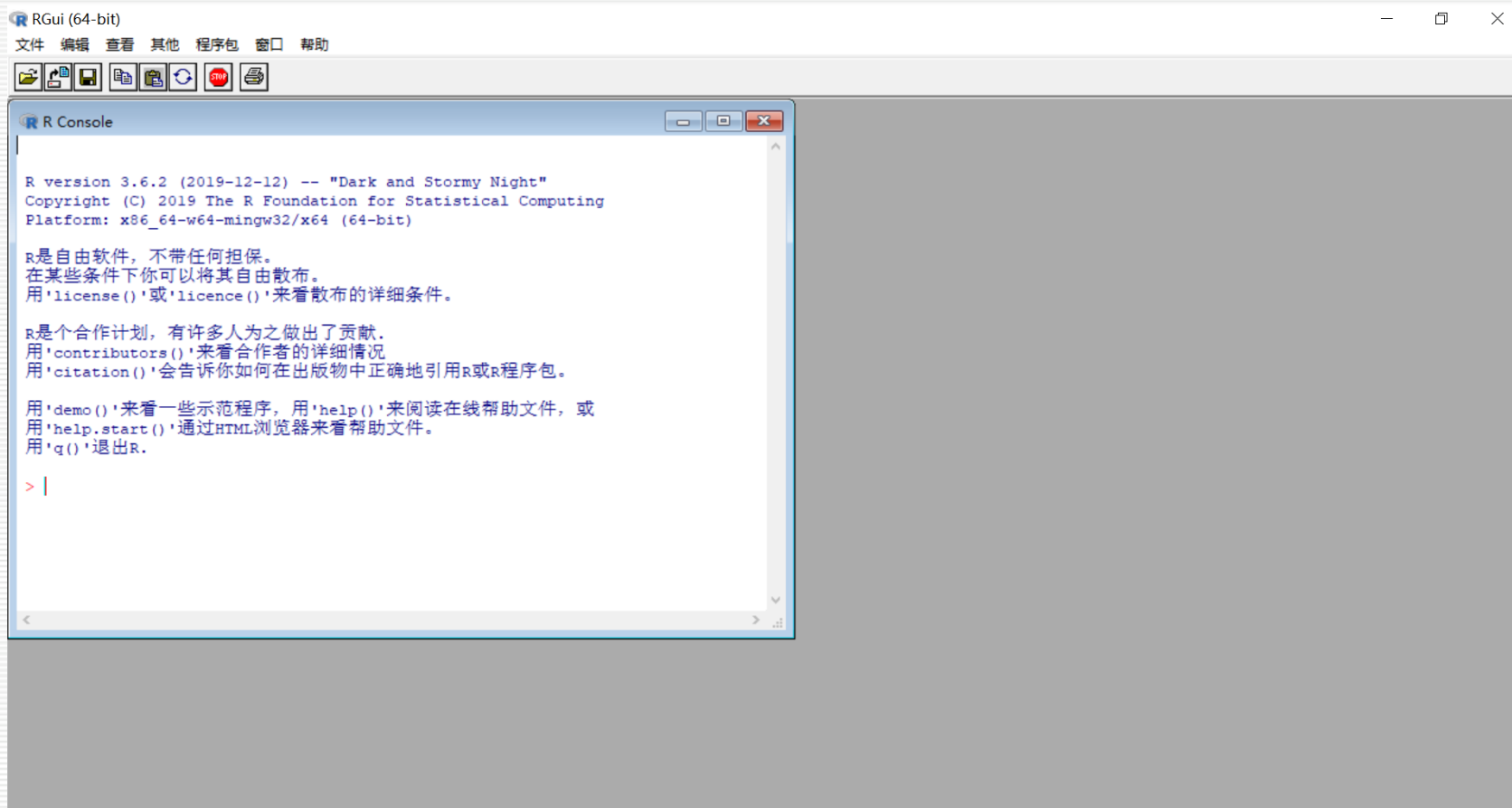
- [Does R run under my version of Windows?](#)
- [How do I update packages in my previous version of R?](#)

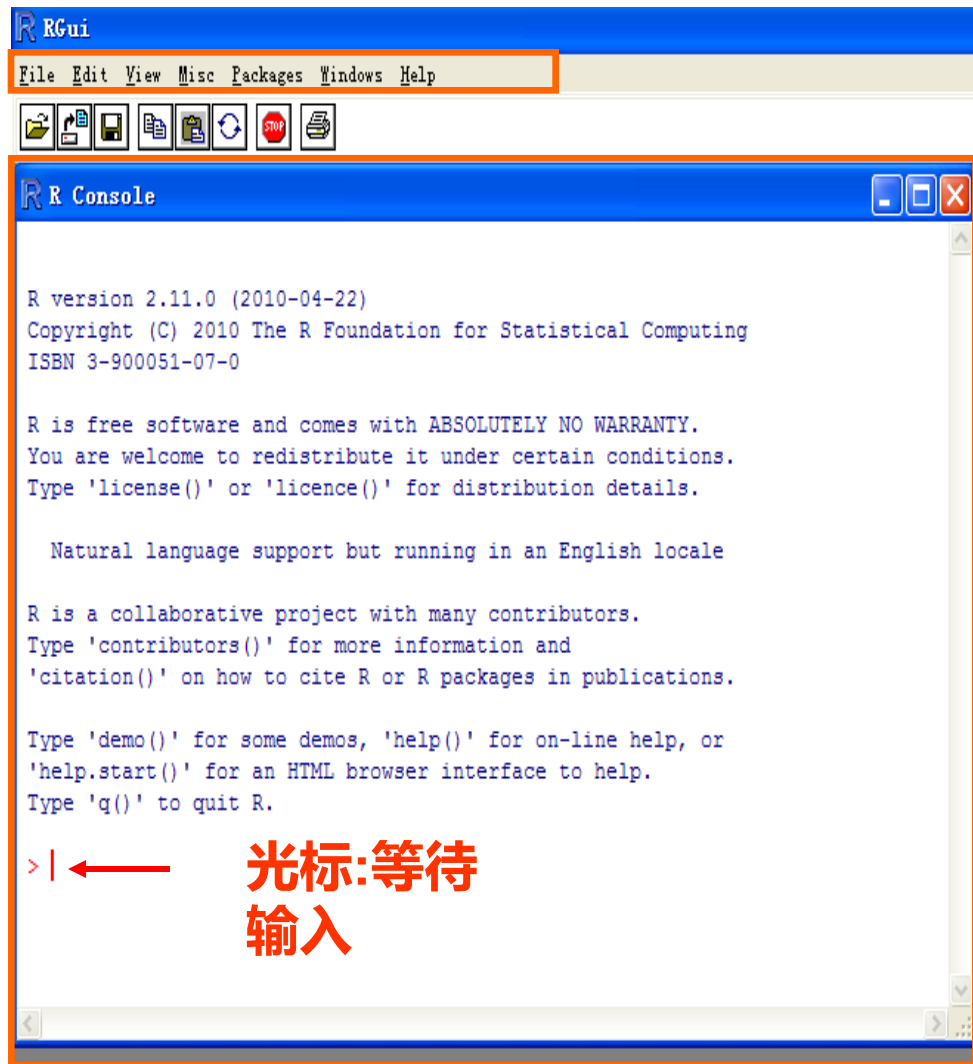
Please see the [R FAQ](#) for general information about R and the [R Windows FAQ](#) for Windows-specific information.

Other builds

- Patches to this release are incorporated in the [r-patched snapshot build](#).
- A build of the development version (which will eventually become the next major release of R) is available in the [r-devel snapshot build](#).
- [Previous releases](#)

Windows下载和安装R



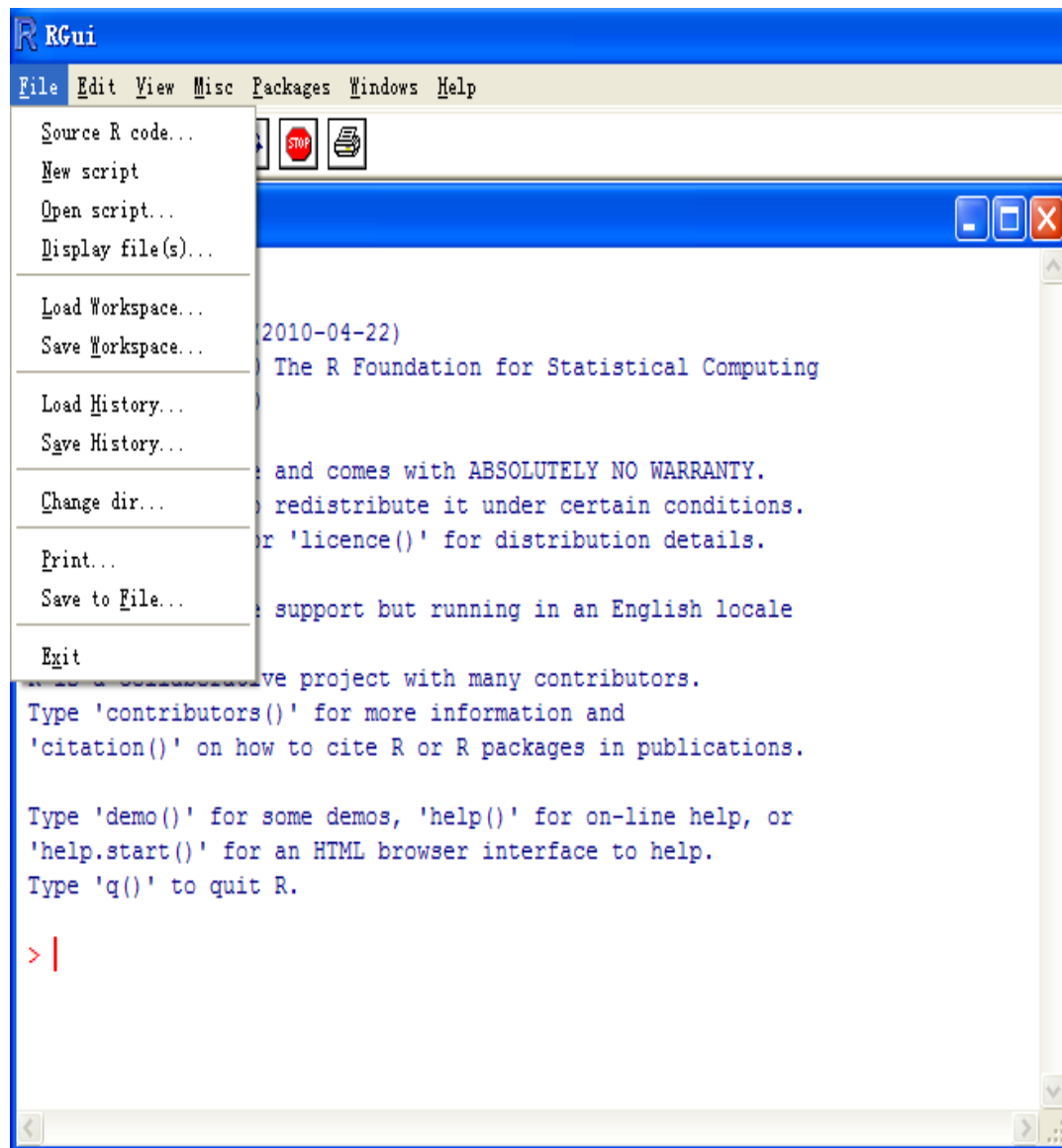


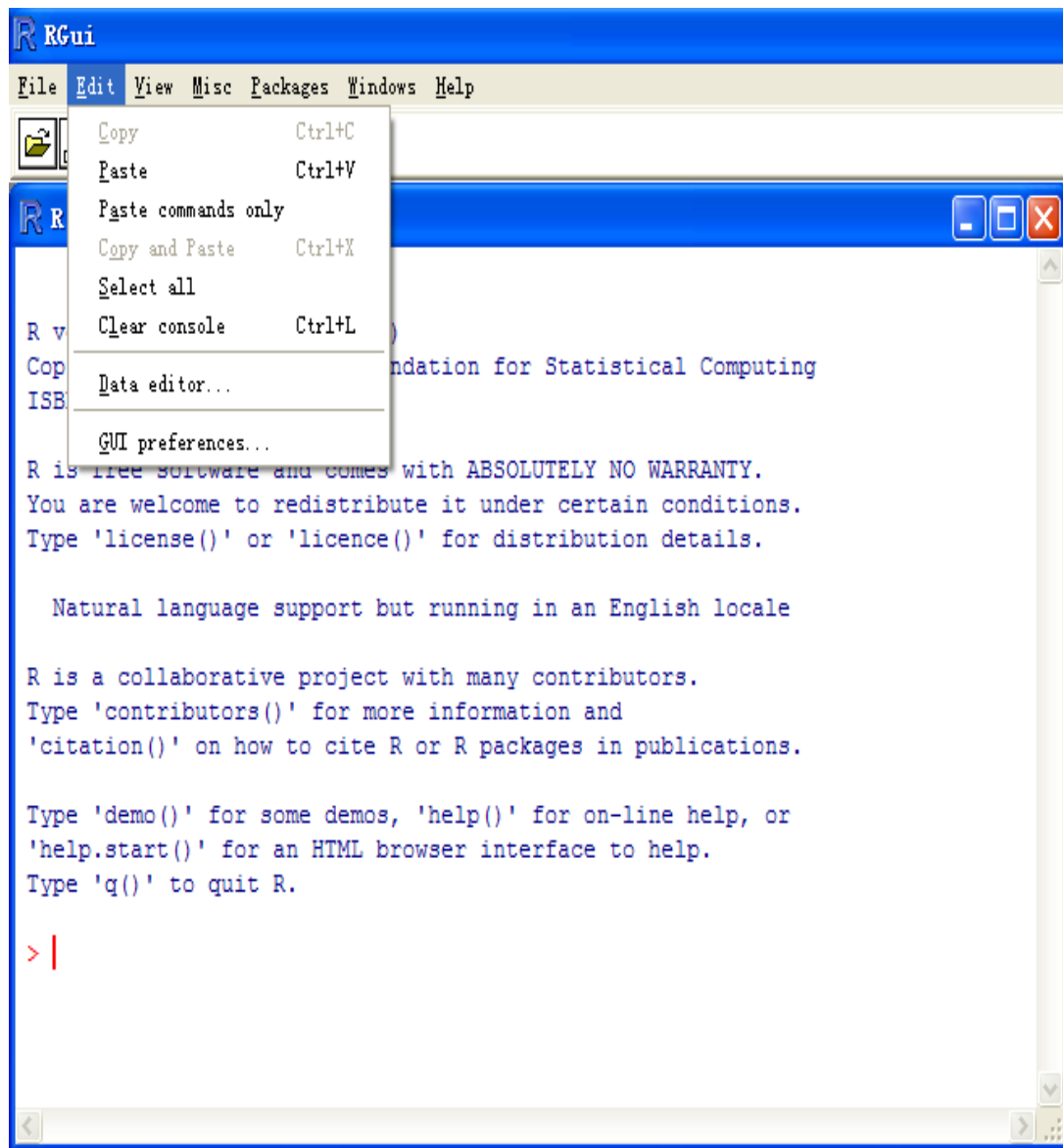
菜单栏
快捷按钮

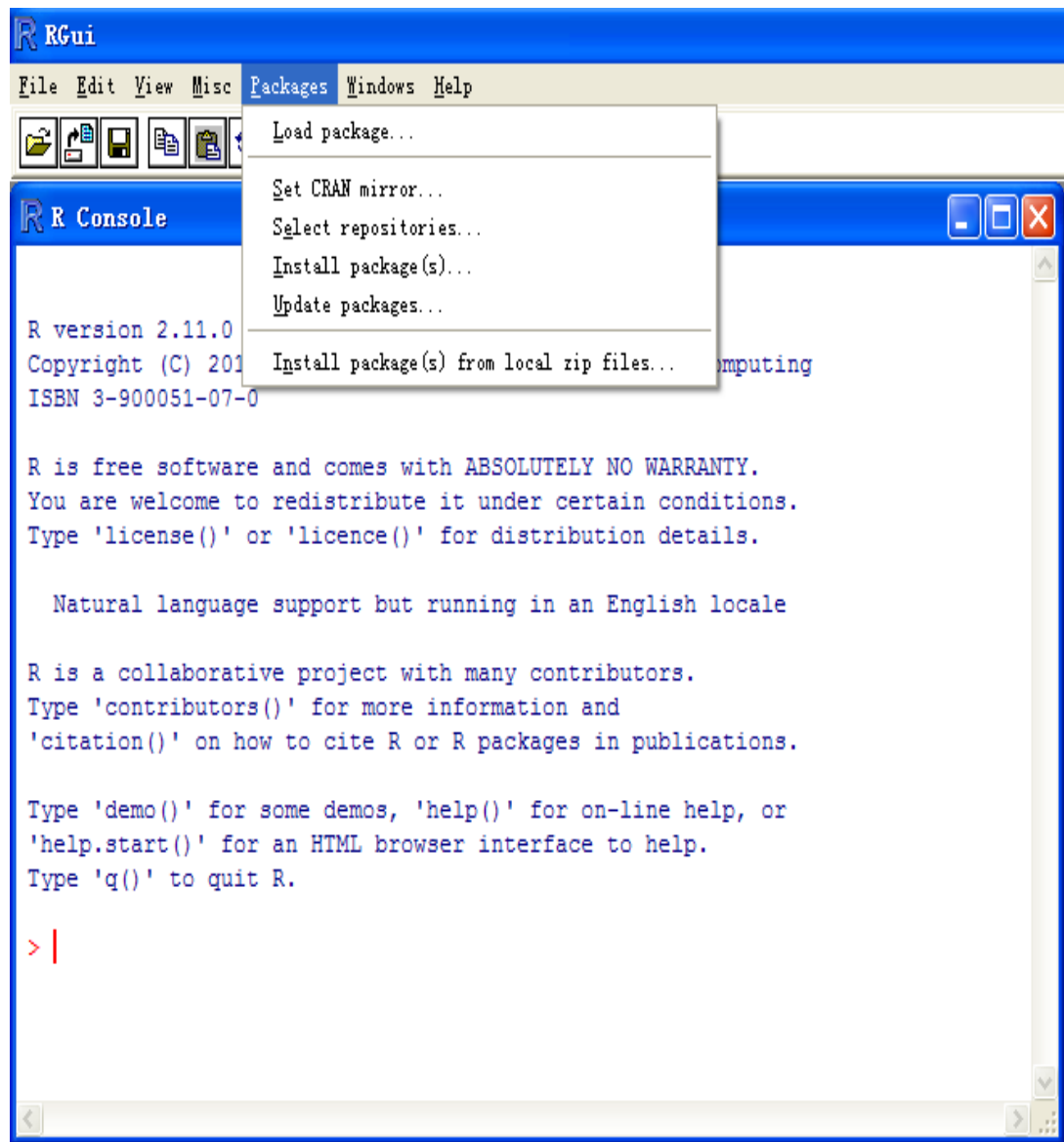
控制台

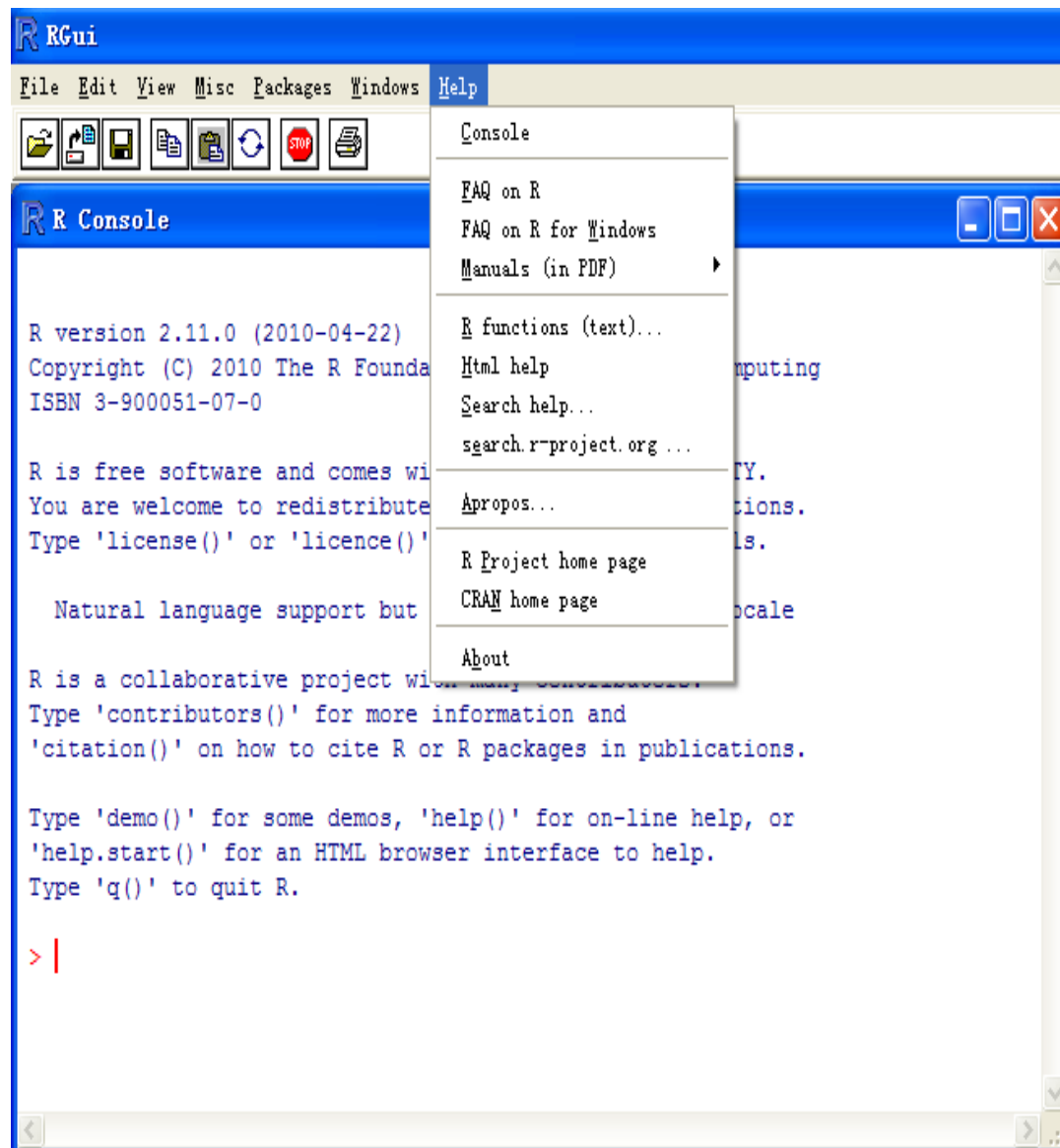
R软件基础界面(Windows版)

RGui + R Console









■ 工具条介绍



- Source R code 执行R文件(*.R或*.r)
- Save image 保存工作空间,文件名为*.RData
- Load image 打开已有的工作空间
- Stop current computation
中止当前计算(由于超时等原因)

Windows下载和安装RStudio

R程序包（R Packages）

程序包是什么？

R程序包是多个函数的集合，具有详细的说明和示例。

Windows环境下的R程序包是经过编译的zip压缩包。

每个程序包包含R函数、数据、帮助文件和描述文件等。

为什么要安装程序包？

R程序包是R功能扩展，特定的分析功能，需要用相应的程序包实现。

例如：系统发育分析常用到ape程序包，群落生态学使用vegan包等。

常用R程序包(I)

ade4	利用欧几里得方法进行生态学数据分析
adephylo	系统进化数据挖掘与比较方法
ape	系统发育与进化分析
apTreeshape	进化树分析
boot	Bootstrap检验
cluster	聚类分析
ecodist	生态学数据相异性分析
FD	功能多样性分析
geiger	物种形成速率与进化分析

常用R程序包(II)

Graphics	绘图
lattice	栅格图
maptools	空间对象的读取和处理
me4a	生态学和生物地理学多元数据处理
mgcv	广义加性模型相关
mvpart	多变量分解
nlme	线性及非线性混合效应模型
ouch	系统发育比较
pgirmess	生态学数据分析
phangorn	系统发育分析

常用R程序包(III)

picante	群落系统发育多样性分析
raster	栅格数据分析与处理
seqinr	DNA序列分析
sp	空间数据处理
spatstat	空间点格局分析，模型拟合与检验
splancs	空间与时空点格局分析
stats	R统计学包
SDMTools	物种分布模型工具
vegan	植物与植物群落的排序，生物多样性计算

vegan: Community Ecology Package

Ordination methods, diversity analysis and other functions for community and vegetation ecologists.

Version: 1.17-2

Suggests: [MASS](#), [mgcv](#), [lattice](#), [cluster](#), [scatterplot3d](#), [rgl](#), tcltk

Published: 2010-03-08

Author: Jari Oksanen, F. Guillaume Blanchet, Roeland Kindt, Pierre Legendre, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens, Helene Wagner

Maintainer: Jari Oksanen <jari.oksanen at oulu.fi>

License: [GPL-2](#)

URL: <http://vegan.r-forge.r-project.org/>

In views: [Environmetrics](#), [Multivariate](#), [Phylogenetics](#), [Psychometrics](#), [Spatial](#)

CRAN
checks: [vegan results](#)

Downloads:

Package source: [vegan_1.17-2.tar.gz](#)

MacOS X binary: [vegan_1.17-2.tgz](#)

Windows binary: [vegan_1.17-2.zip](#)

Reference manual: [vegan.pdf](#)

不同平台上的
的程序包

vegan包页面

R程序包

Downloads:

Reference manual:	BayesFactor.pdf
Vignettes:	Demos and comparisons Vignette menu User's manual Odds and probabilities Prior checks
Package source:	BayesFactor 0.9.12-2.tar.gz
Windows binaries:	r-devel: BayesFactor 0.9.12-2.zip , r-release: BayesFactor 0.9.12-2.zip , r-oldrel: BayesFactor 0.9.12-2.zip
OS X Snow Leopard binaries:	r-release: BayesFactor 0.9.12-2.tgz , r-oldrel: BayesFactor 0.9.11-1.tgz
OS X Mavericks binaries:	r-release: BayesFactor 0.9.12-2.tgz
Old sources:	BayesFactor archive

一般CRAN提供三种版本的安装包：开发的版本(r-devel)、**发行版(r-release)**、以前的版本(r-oldrel)，建议下载安装**发行版**，安全稳定。

安装程序包的方法（重要）

1 直接用R上菜单里面的快捷键选择镜像后安装。

2 用函数 `install.packages()`,

如果已经连接到互联网, 在括号中输入要安装的程序包名称, 选择镜像后, 程序将自动下载并安装程序包。

例如: 要安装中文分词包, 在控制台中输入

```
install.packages("jiebaR")
```

3 安装本地zip包

路径: `Packages>install packages from local files`

选择本地磁盘上存储zip包的文件夹。

程序包使用（重要）

程序包的中函数，都要**先导入**，再使用，因此导入程序包是第一步。

例如，使用分词程序包*jiebaR*时，先在控制台中输入：

```
library(jiebaR)
```

程序包内函数的用法与R内置基本函数的用法一样。

查看程序包帮助文件

程序包内部都有哪些函数？分别有什么功能？

查询程序包内容最常用的方法：

1 菜单 帮助 > html帮助

2 查看pdf帮助文档

3 直接使用help函数，如help(jiebaR)

查看函数的帮助文件

函数的默认值是什么？ 怎么使用？ 使用时需要注意什么问题？

我们需要查询函数的帮助。

1 `?t.test`

2 `Rgui > Help > Html help`

3 `apropos("t.test")` #模糊查找

4 `help("t.test")`

5 `help.search("t.test")`

6 `??"t.test"`

7 查看R包pdf手册

Fitting Linear Models

Description

`lm` is used to fit linear models. It can be used to carry out regression, single stratum analysis of variance and analysis of covariance (although [aov](#) may provide a more convenient interface for these).

Usage

```
lm(formula, data, subset, weights, na.action,  
   method = "qr", model = TRUE, x = FALSE, y = FALSE, qr = TRUE,  
   singular.ok = TRUE, contrasts = NULL, offset, ...)
```

Arguments

<code>formula</code>	an object of class " formula " (or one that can be coerced to that class): a symbolic description of the model to be fitted. The details of model specification are given under ‘Details’.
<code>data</code>	an optional data frame, list or environment (or object coercible by as.data.frame to a data frame) containing the variables in the model. If not found in <code>data</code> , the variables are taken from <code>environment(formula)</code> , typically the environment from which <code>lm</code> is called.
<code>subset</code>	an optional vector specifying a subset of observations to be used in the fitting process.

帮助文件的内容

<code>lm{stats}</code>	#函数名及所在包
<code>Fitting Linear Models</code>	#标题
<code>Description</code>	#函数描述
<code>Usage</code>	#默认选项
<code>Arguments</code>	#参数
<code>Details</code>	#详情
<code>Author(s)</code>	#作者
<code>References</code>	#参考文献
<code>Examples</code>	#举例

基本语法

- R的基本界面是一个交互式命令窗口，命令提示符是一个大于号，命令的结果马上显示在命令下面。
- R命令主要有两种形式：表达式或赋值运算（用<-或=表示）
- 在命令提示符后键入一个表达式，表示计算此表达式并显示计算结果
- 赋值运算把赋值号右边的值计算出来赋给左边的变量
- 可以用向上光标键来找回以前运行的命令，再次运行或修改后再运行
- R是区分大小写的，因此x和X是不同的对象名，对应的对象也不相同

画图练习

我们用一些例子来看R软件的特点。

假设我们已经进入了R的交互式窗口。

如果没有打开的图形窗口，在R中，用：`> x11()`

可以打开一个作图窗口。然后，输入以下语句：

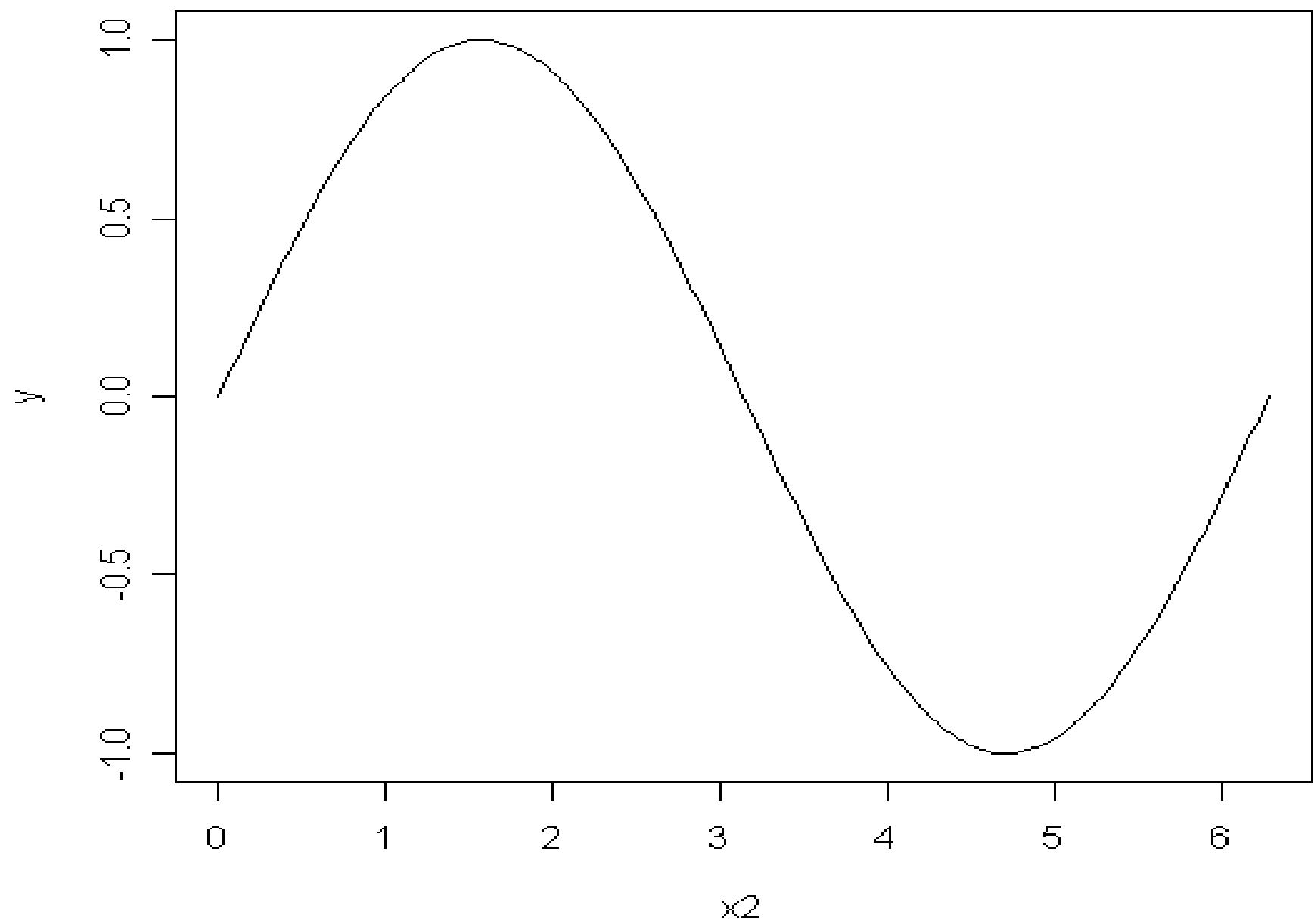
```
> x1 <- 0:100
```

```
> x2 <- x1*2*pi/100
```

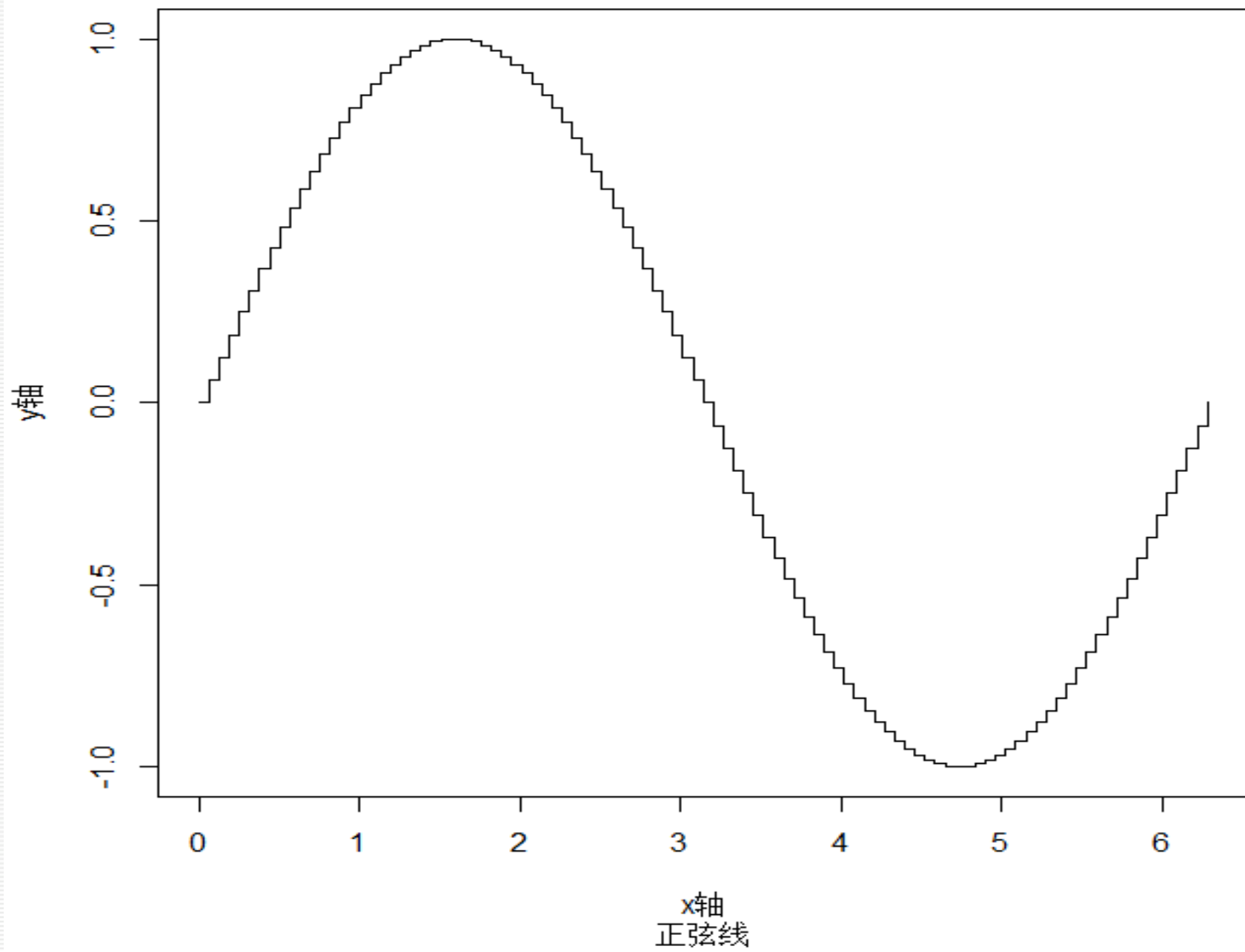
```
> y <- sin(x2)
```

```
> plot(x2,y, type='l')
```

```
> plot(x2,y,main="画图练习", type="s", sub="正弦线",  
xlab="x轴",ylab='y轴')
```



画图练习



画图练习

- 使用这些语句绘制正弦曲线图。其中：“<-”是赋值运算符。0:100表示一个从0到100的等差数列向量。
- 从第二个语句可以看出，我们可以对向量直接进行四则运算。例如，计算得到的x2是向量x1的所有元素乘以常数 $2\pi/100$ 的结果。
- 从第三个语句可看到，向量可以作为函数的输入，并可以输出一个结果向量y，y的每一个分量是自变量x2的每一个分量经过函数（正弦）运算后得到的值。
- 从最后一个语句可以看出，函数的调用也很自由，函数的参数变量可以按位置给出，也可以用“**变量名=**”的形式指定变量值，还可以使用缺省值。

课后练习

1. 下载并安装R软件和R Studio软件
2. 在 $[-1, 2]$ 上画 $y = e^{2x} + \sin(3x^2)$ 的图形
3. “莲花盛开” 画图练习

工作路径与文件读取

- `getwd()`: 获取工作路径
- `setwd()`: 设置工作路径
- `read.csv()`: 读取由逗号分隔(comma-separated values)的文件信息
 - 后缀均为.csv, 只需给出文件名即可读取