

R 语言-1

郑泽靖 zzjstat2023@163.com

北京师范大学统计学院

2025 年 9 月 26 日

计算 10 个数的平均值

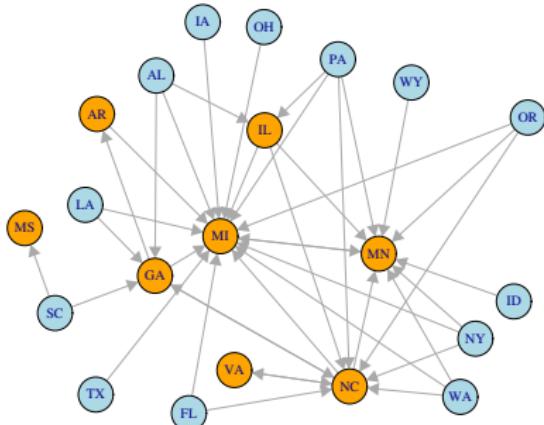
- 示例：数字集合 $\{4, 8, 15, 16, 23, 42, 7, 9, 12, 11\}$
- 计算步骤：
 - ① 求和： $4 + 8 + 15 + 16 + 23 + 42 + 7 + 9 + 12 + 11 = 147$
 - ② 计算平均值：147 除以 10，得到 14.7
- 结论：该集合的平均值为 14.7

手动计算的局限性

- 挑战：
 - 手动分析大规模数据集既繁琐又低效。
 - 为确保准确性和效率，自动化工具显得尤为重要。
- 解决方案：
 - 使用编程语言（如 R）来快速计算和处理数据。
 - 编写简洁的代码来有效地处理大规模数据集。
- 优势：
 - 快速：几秒内完成计算。
 - 准确：减少人为错误。
 - 可扩展：轻松处理更大数据集。

为什么统计需要编程

- 大规模数据处理
- 自动化与重复性
- 复杂计算
- 数据可视化



R 的优点

- **自由软件：**免费、开放源代码，支持各个主要计算机系统；

Download and Install R

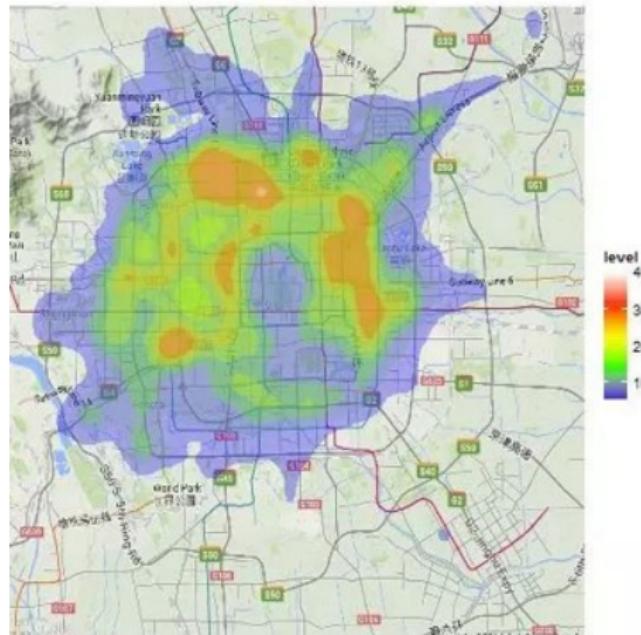
Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- [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.

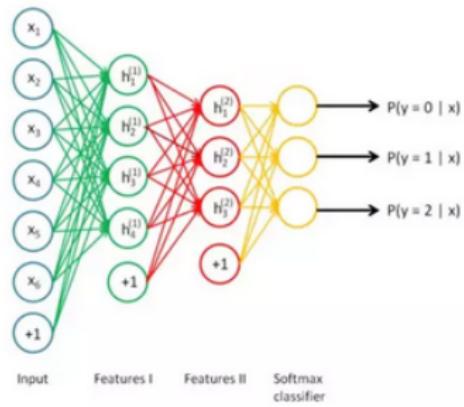
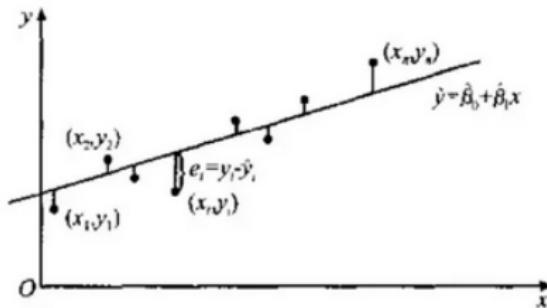
R 的优点

- 强大的图形功能：强调交互式数据分析，支持复杂算法描述；



R 的优点

- 丰富的统计方法：实现了经典和现代的统计方法，包括：
 - 参数和非参数假设检验
 - 线性回归与广义线性回归等



R 的优点

- 广泛的使用和扩展包：统计科研工作者广泛使用 R 进行计算和发表算法。截止 2023 年 4 月，R 扩展包的主要分发网站 CRAN 上已有近 19,000 个扩展包。

All experiments in this paper are conducted in R. The GLM Lasso is implemented via R package `glmnet` (Friedman et al., 2010). We summarize R codes for GLM transfer learning algorithms in a new R package `glmtrans`, which is available on CRAN. We use 10-fold cross-validation to choose the penalty parameter for naïve-Lasso and our GLM transfer learning algorithms. The largest λ which achieves one standard error within the minimum

R 的优点

- 便捷的程序交互：R 语言提供了友好的交互式环境。
- 多种数据源支持：可以从多种数据源导入数据，包括文本文件、数据库等。
- 丰富的社区文化：活跃的社区支持，如 GitHub、CSDN。



R 的下载与安装

步骤 1：访问 R 项目官网，点击“Download R”。



[Home]

Download

CRAN

R Project

About R

Logo

Contributors

What's New?

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.

R 的下载与安装

步骤 2：在弹出的镜像（Mirrors）页面上选择合适的镜像入口。如果你在中国，可以直接选择“China”下离你最近的一个镜像。

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.sjtu.edu.cn/cran/>
<https://mirrors.sustech.edu.cn/CRAN/>
<https://mirrors.nwafu.edu.cn/cran/>

Colombia

Colombia

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)

R 的下载与安装

步骤 3：点击对应的操作系统以进行下载。

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

R 的下载与安装

步骤 4：选择 **base**（基本功能版本），点击“Install R for the first time”。

R for Windows

Subdirectories:

<u>base</u>	Binaries for base distribution. This is what you want to install R for the first time .
<u>contrib</u>	Binaries of contributed CRAN packages (for R >= 3.4.x).
<u>old contrib</u>	Binaries of contributed CRAN packages for outdated versions of R (for R < 3.4.x).
<u>Rtools</u>	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.

R 的下载与安装

步骤 5：下载 R-x.x.x 并进行安装。

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.

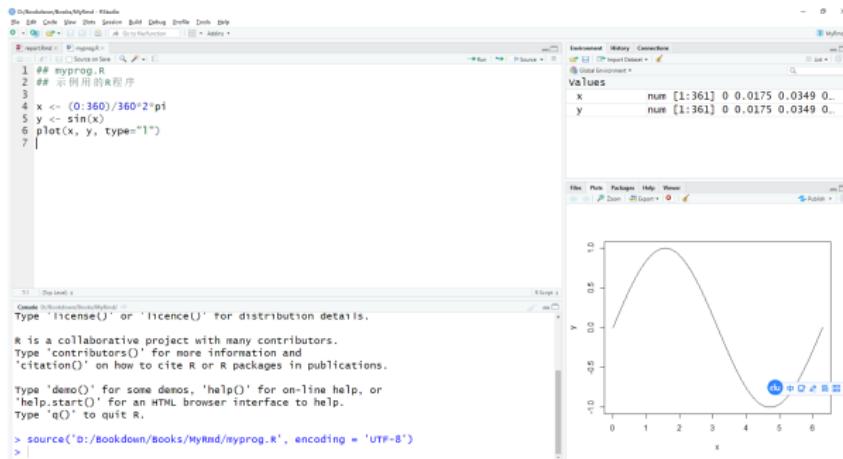
Frequently asked questions

- [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.

RStudio

RStudio 是一个功能更强大的 R 图形界面。安装好 R 的官方版本后，安装 RStudio 可以更方便地使用 R (<https://posit.co/downloads/>)。



The screenshot shows the RStudio interface with the following components:

- Code Editor:** Displays the R script "myprog.R" containing the following code:

```
## myprog.R
## 简例用的R程序
x <- (0:360)/360*2*pi
y <- sin(x)
plot(x, y, type="l")
```
- Console:** Shows the R startup message and the command to source the script:

```
License or 'licence()' for distribution details.  
R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.  
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.  
> source('d:/bookdown/books/MyRmd/myprog.R', encoding = 'UTF-8')
```
- Environment View:** Shows two variables: x and y, both of which are numeric vectors of length 361.
- Plots:** A line plot of the sine wave $y = \sin(x)$ where x ranges from 0 to 2π . The plot includes a legend indicating the data series.

RStudio 的使用

RStudio 界面的主要窗格：

- 编辑窗格：编写和编辑代码。
- 控制台（Console）：直接执行 R 命令。
- 环境（Environment）：显示当前变量和数据。
- 文件（Files）：查看项目文件和数据。
- 图像（Plots）：查看绘制的图像。

如何使用帮助文档：

- 在控制台中使用 `?function_name` 获取帮助。
- 访问 [RDocumentation](#)。

RStudio 的使用：创建文件

创建 R 脚本文件的步骤：

- ① 打开 RStudio。
- ② 点击 File > New File > R Script。
- ③ 在编辑窗格中输入代码。
- ④ 点击 File > Save 或使用快捷键 Ctrl + S，
选择保存位置并命名文件（如 example.R）。

R 代码的运行

在 RStudio 中运行 R 代码的方法：

- 在编辑窗格中输入代码后，使用 Ctrl + Enter 运行选中的代码行。
- 你也可以在控制台中直接输入代码并按下 Enter 键运行。
- 如果要运行选中的或者某行代码，可以点击工具栏中的“Run”按钮。

四则运算

四则运算示例：

```
1 # 计算一个简单的表达式  
2 5+(2.3-1.125)*3.2/1.1+1.23E3  
3 ## [1] 1238.418
```

- 输出结果前的井号（#）表示这是注释，帮助我们理解代码的含义。
- 输出的方括号和数字 1 是提示性标记，方便我们识别结果位置。
- 科学记数法（如 1.23E3）是如此表示： 1.23×10^3 。

数学函数

常用数学函数：

- 平方根、指数、对数：

```
1 sqrt(6.25)      # 计算平方根  
2 ## [1] 2.5  
3 exp(1)          # 计算自然指数  
4 ## [1] 2.718282  
5 # 计算以10为底的对数  
6 log10(10000)  
7 ## [1] 4
```

数学函数

- 取整函数:

```
1 round(1.1234, 2)    # 四舍五入  
2 ## [1] 1.12  
3 floor(1.1234)       # 向下取整  
4 ## [1] 1  
5 ceiling(1.1234)      # 向上取整  
6 ## [1] 2
```

三角函数

常用三角函数：

```
1 pi                      # 圆周率
2 ## [1] 3.141593
3 sin(pi/6)              # 正弦
4 ## [1] 0.5
5 cos(pi/6)              # 余弦
6 ## [1] 0.8660254
7 tan(pi/6)               # 正切
8 ## [1] 0.5773503
```

常量

常量定义：

常量是直接写在程序中的值，包括数值、字符串等。

- 数值型常量：整型、单精度、双精度
 - 示例：123, 123.45
- 字符型常量：使用双引号或单引号包围
 - 示例："Li Ming" 或 'Li Ming'
- 逻辑型常量：只有 TRUE 和 FALSE

常量

- 缺失值: 用 NA 表示, 常见于数据缺失情况。
- NaN (Not a Number): 表示未定义或不可表示的数值, 如 0 除以 0。
- Inf: 表示正无穷或负无穷, 通常在除以零时出现。

变量

变量定义：

变量用于保存输入值或计算结果。在 R 中，变量可以保存各种数据类型，如标量、向量、矩阵、数据框和函数。

- 变量名规则：

- 由字母、数字、下划线和句点组成，不能以数字开头。
- 变量名区分大小写，如 y 和 Y 是不同的变量。
- 示例：x, x1, X, cancer.tab

- 赋值：使用 <- 或 = 定义变量。

作业要求

请在 R 中执行文件中的代码，并观察结果。将结果截图或粘贴到 Word 文档中，最后将文档保存为 Word 或 PDF 格式提交。

- 请确保在执行代码时，能够观察到每个变量的值和计算结果。
- 将结果截图或粘贴到 Word 文档中，并确保代码及其输出是清晰可见的。
- 最后，10 月 10 号前提交到微信群里二维码收件箱。主题：作业 1+ 学号 + 姓名，附件为 word 或 pdf。

Thanks!