



R 基础培训

第三届中国 R 语言会议

导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

刘思喆
第三届中国 R 语言会议
北京



导言

R 简介

- 初识 R
- 下载和安装
- 帮助系统
- 图形界面
- 包

数据处理

- 输入输出
- 数据结构
- 数据操作及变换
- 日期
- 字符处理
- 高级数据处理

函数及运算

- 数学函数
- 矩阵运算

编程

绘图

- 基本元素
- 图形设备

统计

- 分布与抽样
- 假设检验
- 拿“回归”来说事

有用的技巧

Ross Ihaka and Robert Gentleman



What is R(I)



R HomePage

R is 'GNU S', a freely available language and environment for statistical computing and graphics which provides a wide variety of statistical and graphical techniques: linear and nonlinear modelling, statistical tests, time series analysis, classification, clustering, etc.

Our introduction to the R environment did not mention statistics, yet many people use R as a statistics system.

We prefer to think of it of an environment within which many classical and modern statistical techniques have been implemented.

From R introduction

导言

R 简介

- 初识 R
- 下载和安装
- 帮助系统
- 图形界面
- 包

数据处理

- 输入输出
- 数据结构
- 数据操作及变换
- 日期
- 字符处理
- 高级数据处理

函数及运算

- 数学函数
- 矩阵运算

编程

绘图

- 基本元素
- 图形设备

统计

- 分布与抽样
- 假设检验
- 拿“回归”来说事

有用的技巧

What is R(II)



导言

R 简介

- 初识 R
- 下载和安装
- 帮助系统
- 图形界面
- 包

数据处理

- 输入输出
- 数据结构
- 数据操作及变换
- 日期
- 字符处理
- 高级数据处理

函数及运算

- 数学函数
- 矩阵运算

编程

绘图

- 基本元素
- 图形设备

统计

- 分布与抽样
- 假设检验
- 拿“回归”来说事

有用的技巧

- ① R is data analysis software
- ② R is a programming language
- ③ R is an environment for statistical analysis
- ④ R is an open-source software project
- ⑤ R is a community



目录

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

SAS、SPSS、R 的比较:

主题	SAS Product	SPSS Product	R Package
高级模块	SAS/STAT®	SPSS Advanced Models™	stat, MASS
基础模块	SAS®	SPSS base™	base
表格	SAS Base®Report、SQL、 Tabulate	SPSS Custom Tables™	reshape, sqldf
关联分析	SAT/STAT®:Correp	SPSS Categories™	MASS, ade4, cocorresp, ...
数据接口	SAS/ACCESS®	SPSS Data Access Pack™	DBI, foreign, RODBC
矩阵运算	SAS/IML®, SAS/IML Work- shop®	SPSS Matrix™	matlab, Matrix, sparseM
数据挖掘	Enterprise Miner™	Clementine®	rattle, arules, FactoMineR
GIS	SAS/GIS®, SAS/Graph®	SPSS Map™	maps, mapdata, mapproj, ...
时间序列	SAS/ETS®	SPSS Trends™Expert Mod- eler	over 40 packages
结构方程	SAS/STAT®: Calis	Amos™	sem
决策树	Enterprise Miner™	SPSS Classification Trees™, AnswerTree™	ada, adabag, BayesTree, boost, party, rpart, tree, ...

- ACM Software System Award(1998)
- John Chambers
- The Statistical Computing and Graphics Award(2010,BI-annual)



导言

R 简介

初识 R
下载和安装
帮助系统
图形界面
包

数据处理

输入输出
数据结构
数据操作及变换
日期
字符处理
高级数据处理

函数及运算

数学函数
矩阵运算

编程

绘图

基本元素
图形设备

统计

分布与抽样
假设检验
拿“回归”来说事

有用的技巧



导言

R 简介

- 初识 R
- 下载和安装
- 帮助系统
- 图形界面
- 包

数据处理

- 输入输出
- 数据结构
- 数据操作及变换
- 日期
- 字符处理
- 高级数据处理

函数及运算

- 数学函数
- 矩阵运算

编程

绘图

- 基本元素
- 图形设备

统计

- 分布与抽样
- 假设检验
- 拿“回归”来说事

有用的技巧

- ACM Software System Award(1998)
- John Chambers
- The Statistical Computing and Graphics Award(2010,BI-annual)

ACM Software System Award:

- 1983 Unix
- 1986 TeX
- 1989 PostScript
- 1991 TCP/IP
- 1995 World-Wid-Web
- 1997 Tcl/Tk
- 1998 S
- 1999 The Apache Group
- 2002 Java



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

R 不是万能的：

- 内存
- 多线程



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

- ❶ is not a database, but connects to DBMSs
- ❷ language interpreter can be very slow, but allows to call own C/C++ code
- ❸ no spreadsheet view of data, but connects to Excel
- ❹ has no graphical user interfaces, but connects to Java, TclTk

R 在 Windows 下的操作平台

R 基础培训



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

```
R R Console
File Edit Misc Packages Windows Help

R version 2.11.0 (2010-04-22)
Copyright (C) 2010 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type '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.

> |
```



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

← → ↻ ⓘ ☆ <https://www.r-project.org/>



About R
[What is R?](#)
[Contributors](#)
[Screenshots](#)
[What's new?](#)

Download, Packages
[CRAN](#)

R Project
[Foundation](#)
[Members & Donors](#)
[Mailing Lists](#)
[Bug Tracking](#)
[Developer Page](#)
[Conferences](#)
[Search](#)

Documentation
[Manuals](#)
[FAQs](#)
[The R Journal](#)
[Wiki](#)
[Books](#)
[Certification](#)
[Other](#)

Misc
[Bioconductor](#)
[Related Projects](#)
[User Groups](#)
[Links](#)

CRAN Mirrors

The Comprehensive R Archive Network is available at the following URLs, please choose a location close to you. Some statistics [here](#).

Australia
<http://cran.ms.unimelb.edu.au/>

University of Melbourne

Austria
<http://cran.at.r-project.org/>

Wirtschaftsuniversitaet Wien

Belgium
<http://www.freeststatistics.org/cran/>

K.U.Leuven Association

Brazil
<http://cran-r.c3sl.ufpr.br/>
<http://cran.fiocruz.br/>
<http://www.vps.fmvz.usp.br/CRAN/>
<http://brieger.esalq.usp.br/CRAN/>

Universidade Federal do Parana
Oswaldo Cruz Foundation, Rio de Janeiro
University of Sao Paulo, Sao Paulo
University of Sao Paulo, Piracicaba

Canada
<http://cran.stat.sfu.ca/>
<http://mirror.its.dal.ca/cran/>
<http://probability.ca/cran/>
<http://cran.skazkaforyou.com/>
<http://cran.parentinginformed.com/>

Simon Fraser University, Burnaby
Dalhousie University, Halifax
University of Toronto
iWeb, Montreal
iWeb, Montreal

Chile
<http://dirichlet.mat.puc.cl/>

Pontificia Universidad Catolica de Chile, Santiago

China
<http://ftp.ctex.org/mirrors/CRAN/>
<http://cran.csdn.cn/>
<http://mirrors.geoxpat.com/cran/>

CTEX.ORG
Computer Network Information Center, CAS, Beijing
GeoXpat.Com

Colombia
<http://www.laqee.unal.edu.co/CRAN/>

National University of Colombia

Denmark
<http://mirrors.dotsrc.org/cran/>

dotsrc.org, Aalborg

France

[导言](#)[R 简介](#)[初识 R](#)[下载和安装](#)[帮助系统](#)[图形界面](#)[包](#)[数据处理](#)[输入输出](#)[数据结构](#)[数据操作及变换](#)[日期](#)[字符处理](#)[高级数据处理](#)[函数及运算](#)[数学函数](#)[矩阵运算](#)[编程](#)[绘图](#)[基本元素](#)[图形设备](#)[统计](#)[分布与抽样](#)[假设检验](#)[拿“回归”来说事](#)[有用的技巧](#)

- ① Linux
- ② MacOS X
- ③ Windows

R 大概每 3 个月更新一次版本：

```
update.packages(checkBuilt = TRUE, ask = FALSE)
# or install.packages(.packages(all = TRUE))
```

1

- `help(topic)`
- `?topic`
- `help.search('topic')`
- `apropos('topic')`
- `help.start()`
- `RSiteSearch('topic')`



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

?topic 的内容

- 1 Description — 函数描述
- 2 Usage — 默认选项
- 3 Arguments — 参数
- 4 Details — 详情
- 5 Value — 数值说明
- 6 Warning — 警告
- 7 Note — 注释
- 8 Author(s) — 作者
- 9 References — 参考文献
- 10 See Also — 参考其他
- 11 Examples — 举例

demo() 和 example()



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

- <http://www.r-project.org/mail.html>
- <http://journal.r-project.org/>



CRAN Task Views

Bayesian	Bayesian Inference
ChemPhys	Chemometrics and Computational Physics
ClinicalTrials	Design, Monitoring, and Analysis of Clinical Trials
Cluster	Cluster Analysis & Finite Mixture Models
Distributions	Probability Distributions
Econometrics	Computational Econometrics
Environmetrics	Analysis of Ecological and Environmental Data
ExperimentalDesign	Design of Experiments (DoE) & Analysis of Experimental Data
Finance	Empirical Finance
Genetics	Statistical Genetics
Graphics	Graphic Displays & Dynamic Graphics & Graphic Devices & Visualization
gR	gRaphical Models in R
HighPerformanceComputing	High-Performance and Parallel Computing with R
MachineLearning	Machine Learning & Statistical Learning
MedicalImaging	Medical Image Analysis
Multivariate	Multivariate Statistics
NaturalLanguageProcessing	Natural Language Processing
Optimization	Optimization and Mathematical Programming
Pharmacokinetics	Analysis of Pharmacokinetic Data
Phylogenetics	Phylogenetics, Especially Comparative Methods
Psychometrics	Psychometric Models and Methods
Robust	Robust Statistical Methods
SocialSciences	Statistics for the Social Sciences
Spatial	Analysis of Spatial Data
Survival	Survival Analysis
TimeSeries	Time Series Analysis

导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

Rcmdr

Commander ()

rattle

rattle()



[引言](#)

[R 简介](#)

[初识 R](#)

[下载和安装](#)

[帮助系统](#)

[图形界面](#)

[包](#)

[数据处理](#)

[输入输出](#)

[数据结构](#)

[数据操作及变换](#)

[日期](#)

[字符处理](#)

[高级数据处理](#)

[函数及运算](#)

[数学函数](#)

[矩阵运算](#)

[编程](#)

[绘图](#)

[基本元素](#)

[图形设备](#)

[统计](#)

[分布与抽样](#)

[假设检验](#)

[拿“回归”来说事](#)

[有用的技巧](#)

初始加载的包

search() 函数返回当前加载的包的情况或者获得本地安装的包列表：

```
> .packages(all.available = TRUE)
[1] "adabag"          "adapt"           "akima"
[4] "amap"            "animation"       "aplpack"
[7] "arules"          "base"            "BayesTree"
[10] "biglm"           "bitops"          "bnlearn"
```

自动加载若干 Packages:

包	描述
stats	常用统计函数
graphics	基础绘图函数
grDevices	基础或 grid 图形设备
utils	R 工具函数
datasets	基础数据集
methods	用于 R 对象和编程工具的方法和类的定义
base	基础函数



导言

R 简介

- 初识 R
- 下载和安装
- 帮助系统
- 图形界面

包

数据处理

- 输入输出
- 数据结构
- 数据操作及变换
- 日期
- 字符处理
- 高级数据处理

函数及运算

- 数学函数
- 矩阵运算

编程

绘图

- 基本元素
- 图形设备

统计

- 分布与抽样
- 假设检验
- 拿“回归”来说事

有用的技巧

Available Packages

Currently, the CRAN package repository features 2420 available packages.

[A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#)

ADGofTest	Anderson-Darling GoF test
ADaCGH	Analysis of data from aCGH experiments
AER	Applied Econometrics with R
AGSDest	Estimation in adaptive group sequential trials
AICcmodavg	Model selection and multimodel inference based on (Q)AIC(c)
AIGIS	Areal Interpolation for GIS data
AIM	AIM: adaptive index model
AIS	Tools to look at the data ("Ad Inidicia Spectata")
ALS	multivariate curve resolution alternating least squares (MCR-ALS)
AMORE	A MORE flexible neural network package
AcceptanceSampling	Creation and evaluation of Acceptance Sampling Plans
AdMitt	Adaptive Mixture of Student-t distributions
AdaptFit	Adaptive Semiparametric Regression
AlgDesign	Algorithmic Experimental Design
Amelia	Amelia II: A Program for Missing Data
AnalyzefMRI	Functions for analysis of fMRI datasets stored in the ANALYZE or NIFTI format
Animal	Analyze time-coded animal behavior data
AquaEnv	AquaEnv - an integrated development toolbox for aquatic chemical model generation

source()



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

- =
- < -, - >
- << -, - >>



导言

R 简介

- 初识 R
- 下载和安装
- 帮助系统
- 图形界面
- 包

数据处理

输入输出

- 数据结构
- 数据操作及变换
- 日期
- 字符处理
- 高级数据处理

函数及运算

- 数学函数
- 矩阵运算

编程

绘图

- 基本元素
- 图形设备

统计

- 分布与抽样
- 假设检验
- 拿“回归”来说事

有用的技巧



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

在开始使用 R 工作前，设置 R 的工作空间 (working directory)：使用 **getwd()** 和 **setwd()** 来获得和设置工作目录。

Listing 1: working directory

```
> getwd()
[1] "C:/Documents and Settings/T61"
> setwd("D:/doc/DataMR/")
> dir()
[1] "1223344866467.zip" "article"   "Liu_article.rar"
[5] "slide1"    "slide2"
```

1

3

5



导言

R 简介

初识 R
下载和安装
帮助系统
图形界面
包

数据处理

输入输出
数据结构
数据操作及变换
日期
字符处理
高级数据处理

函数及运算

数学函数
矩阵运算

编程

绘图

基本元素
图形设备

统计

分布与抽样
假设检验
拿“回归”来说事

有用的技巧

读取数据 I

常用读取数据的函数:

`read.table`、`read.csv`、`read.delim`、`read.fwf`

`card.csv`:

	card	ty	time
1	10001	sale	1
2	10002	return	2
3	10003	sale	3
4	10004	return	4
5	10002	sale	5
6	10003	return	6

Example:

```
> x <- read.csv('card.csv', header = TRUE)
> x
      card      ty time
1 10001    sale    1
2 10002  return    2
3 10003    sale    3
4 10004  return    4
5 10002    sale    5
6 10003  return    6
```

2

4

6

8



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

- **data.entry(x)** 电子数据表形式的录入编辑器
- **scan(x)** 从控制台或文件中读取数据为向量或列表
- **data(x)** 加载指定的数据
- **edit()** 调用文本编辑器修改 R 对象
- **fix(x)** ‘fix’ 调用 ‘edit’ 修改 ‘x’



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

Package:foreign

Functions for reading and writing data stored by statistical packages such as Minitab, S, SAS, SPSS, Stata, Systat, ..., and for reading and writing dBase files.



导言

R 简介

- 初识 R
- 下载和安装
- 帮助系统
- 图形界面
- 包

数据处理

输入输出

- 数据结构
- 数据操作及变换
- 日期
- 字符处理
- 高级数据处理

函数及运算

- 数学函数
- 矩阵运算

编程

绘图

- 基本元素
- 图形设备

统计

- 分布与抽样
- 假设检验
- 拿“回归”来说事

有用的技巧

`write.table`、`write.csv`

Tips:

```
> write.csv(x , 'D:/data.csv' , row.names = FALSE)
> write.table(x,"clipboard",sep = "\t",col.names=NA)
# For Excel
```

1

3



导言

R 简介

- 初识 R
- 下载和安装
- 帮助系统
- 图形界面
- 包

数据处理

输入输出

- 数据结构
- 数据操作及变换
- 日期
- 字符处理
- 高级数据处理

函数及运算

- 数学函数
- 矩阵运算

编程

绘图

- 基本元素
- 图形设备

统计

- 分布与抽样
- 假设检验
- 拿“回归”来说事

有用的技巧

```
> m <- NA
> class(NA)
[1] "logical"
> class(T)
[1] "logical"
> is.na(m)
[1] TRUE
```

- ❶ **na.omit(x)** 去除缺失值 (NA), 如果 x 为矩阵或数据框, 去除相关行
- ❷ **na.fail(x)** 返回错误信息如果 x 包含至少一个 NA



导言

R 简介

- 初识 R
- 下载和安装
- 帮助系统
- 图形界面
- 包

数据处理

- 输入输出
- 数据结构
- 数据操作及变换
- 日期
- 字符处理
- 高级数据处理

函数及运算

- 数学函数
- 矩阵运算

编程

绘图

- 基本元素
- 图形设备

统计

- 分布与抽样
- 假设检验
- 拿“回归”来说事

有用的技巧

from:to 产生一个序列;
需要注意的是 “:” 有较高的优先级, 比如:

```
> 1:10 + 2
[1] 3 4 5 6 7 8 9 10 11 12
> 1:(10 + 2)
[1] 1 2 3 4 5 6 7 8 9 10 11 12
```

而更加一般的序列可以通过 **seq(from, to)** 来产生:

```
> seq(1,5,by = 0.5)
[1] 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0
> seq(1,5,length = 11)
[1] 1.0 1.4 1.8 2.2 2.6 3.0 3.4 3.8 4.2 4.6 5.0
```

Objects, their modes and attributes



导言

R 简介

- 初识 R
- 下载和安装
- 帮助系统
- 图形界面
- 包

数据处理

输入输出

数据结构

- 数据操作及变换
- 日期
- 字符处理
- 高级数据处理

函数及运算

- 数学函数
- 矩阵运算

编程

绘图

- 基本元素
- 图形设备

统计

- 分布与抽样
- 假设检验
- 拿“回归”来说事

有用的技巧

```
> length(x)
[1] 3
> class(x)
[1] "data.frame"
> mode(x)
[1] "list"
> attributes(x)
$names
[1] "card" "ty"    "time"

$row.names
[1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

$class
[1] "data.frame"
```



导言

R 简介

初识 R
下载和安装
帮助系统
图形界面
包

数据处理

输入输出

数据结构

数据操作及变换
日期
字符处理
高级数据处理

函数及运算

数学函数
矩阵运算

编程

绘图

基本元素
图形设备

统计

分布与抽样
假设检验
拿“回归”来说事

有用的技巧

	类型	说明
1	字符 (character)	它们常常被引号包围
2	数字 (numeric)	实数向量
3	整数 (integer)	整数向量
4	逻辑 (logical)	逻辑向量 (TRUE=T、FALSE=F)
5	复数 (complex)	复数
6	列表 (list)	S 对象的向量
7	因子 (factor)	常用于标记样本

在 R (S) 语言中，有一点要牢记：

Everything in S is an object;
Every object in S has a class.

[导言](#)[R 简介](#)[初识 R](#)[下载和安装](#)[帮助系统](#)[图形界面](#)[包](#)[数据处理](#)[输入输出](#)[数据结构](#)[数据操作及变换](#)[日期](#)[字符处理](#)[高级数据处理](#)[函数及运算](#)[数学函数](#)[矩阵运算](#)[编程](#)[绘图](#)[基本元素](#)[图形设备](#)[统计](#)[分布与抽样](#)[假设检验](#)[拿“回归”来说事](#)[有用的技巧](#)

data frame(数据框)

可以理解是一个松散的数据集。它可以是由不同类型的列（数字、因子、字符等）组成的类矩阵（matrix-like）。

- **attach**
- **detach**



导言

R 简介

初识 R
下载和安装
帮助系统
图形界面
包

数据处理

输入输出
数据结构

数据操作及变换

日期
字符处理
高级数据处理

函数及运算

数学函数
矩阵运算

编程

绘图

基本元素
图形设备

统计

分布与抽样
假设检验
拿“回归”来说事

有用的技巧

```
> methods(as)
[1] as.array                as.array.default
[3] as.call                 as.character
[5] as.character.condition  as.character.Date
[7] as.character.default    as.character.error
[9] as.character.factor     as.character.hexmode
[11] as.character.numeric_version as.character.octmode

[115] as.single                as.single.default
[117] as.stepfun              as.symbol
[119] as.table                 as.table.default
[121] as.tclObj               as.ts
[123] as.vector               as.vector.factor
```

[导言](#)[R 简介](#)[初识 R](#)[下载和安装](#)[帮助系统](#)[图形界面](#)[包](#)[数据处理](#)[输入输出](#)[数据结构](#)[数据操作及变换](#)[日期](#)[字符处理](#)[高级数据处理](#)[函数及运算](#)[数学函数](#)[矩阵运算](#)[编程](#)[绘图](#)[基本元素](#)[图形设备](#)[统计](#)[分布与抽样](#)[假设检验](#)[拿“回归”来说事](#)[有用的技巧](#)

向量索引

```
x[n]  
x[-n]  
x[1:n]  
x[-(1:n)]  
x[c(1,4,2)]  
x["name"]  
x[x > 3]  
x[x > 3 & x < 5]  
x[x %in% c("a", "and", "the")]
```

第 n 个元素
除了第 n 个元素的 x
前 n 个元素
第 $n+1$ 至最后的元素
指定元素
名为 "name" 的元素
所有大于 3 的元素
区间 (3,5) 的元素
给定组中的元素



列表索引

<code>x[n]</code>	列表显示元素 <code>n</code>
<code>x[[n]]</code>	列表的第 <code>n</code> 个元素
<code>x[["name"]]</code>	名为 "name" 的元素
<code>x\$name</code>	同上

矩阵索引

<code>x[i, j]</code>	下标为 (i,j) 的元素
<code>x[i,]</code>	第 <code>i</code> 行
<code>x[, j]</code>	第 <code>j</code> 列
<code>x[, c(1, 3)]</code>	第 1 和 3 列
<code>x["name",]</code>	名为 "name" 的行

数据框索引 (矩阵索引加下标)

<code>x[["name"]]</code>	列名为 "name" 的列
<code>x\$name</code>	同上

[前言](#)[R 简介](#)[初识 R](#)[下载和安装](#)[帮助系统](#)[图形界面](#)[包](#)[数据处理](#)[输入输出](#)[数据结构](#)[数据操作及变换](#)[日期](#)[字符处理](#)[高级数据处理](#)[函数及运算](#)[数学函数](#)[矩阵运算](#)[编程](#)[绘图](#)[基本元素](#)[图形设备](#)[统计](#)[分布与抽样](#)[假设检验](#)[拿“回归”来说事](#)[有用的技巧](#)



导言

R 简介

初识 R
下载和安装
帮助系统
图形界面
包

数据处理

输入输出
数据结构

数据操作及变换

日期
字符处理
高级数据处理

函数及运算

数学函数
矩阵运算

编程

绘图

基本元素
图形设备

统计

分布与抽样
假设检验
拿“回归”来说事

有用的技巧

- `which.max(x)`
- `which.min(x)`
- `rev(x)`
- `rle(x)`
- `sort(x)`
- `match(x, y)`
- `which(x == a)`



导言

R 简介

初识 R
下载和安装
帮助系统
图形界面
包

数据处理

输入输出
数据结构

数据操作及变换

日期
字符处理
高级数据处理

函数及运算

数学函数
矩阵运算

编程

绘图

基本元素
图形设备

统计

分布与抽样
假设检验
拿“回归”来说事

有用的技巧

cut(x, breaks)

```
> set.seed(10)
> x <- rnorm(10, 0, 100)
> cut(x, breaks = c(-Inf, -100, -50, 0, 50, 100, Inf))
[1] (0,50]      (-50,0]      (-Inf,-100] (-100,-50]
[5] (0,50]      (0,50]      (-Inf,-100] (-50,0]
[9] (-Inf,-100] (-50,0]
6 Levels: (-Inf,-100] (-100,-50] (-50,0] ... (100,Inf]
> x
[1] 1.874617 -18.425254 -137.133055 -59.916772
[5] 29.454513 38.979430 -120.807618 -36.367602
[9] -162.667268 -25.647839
```



```

> m <- NULL
> for( i in 1:2)
+   m[[i]] <- data.frame(matrix(rnorm(4),2,2))
> m
[[1]]
      X1      X2
1 0.6003370 1.0410788
2 0.1481331 -0.6028166

[[2]]
      X1      X2
1 -0.4325380 0.4349807
2  0.8317534 -0.1724106

> do.call(cbind,m)
      X1      X2      X1      X2
1 0.6003370 1.0410788 -0.4325380 0.4349807
2 0.1481331 -0.6028166  0.8317534 -0.1724106

```

1

导言

3

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

7

数据处理

输入输出

数据结构

9

数据操作及变换

日期

字符处理

高级数据处理

11

13

函数及运算

数学函数

矩阵运算

15

编程

17

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

unique(x)、duplicated(x)

```
> x <- c(2000,2001,2002,2000,2004)
> unique(x)
[1] 2000 2001 2002 2004
> duplicated(x)
[1] FALSE FALSE FALSE TRUE FALSE
```

2

4



导言

R 简介

初识 R
下载和安装
帮助系统
图形界面
包

数据处理

输入输出
数据结构
数据操作及变换

日期

字符处理
高级数据处理

函数及运算

数学函数
矩阵运算

编程

绘图

基本元素
图形设备

统计

分布与抽样
假设检验
拿“回归”来说事

有用的技巧

as.Date(), as.POSIXct()

```
d1 <- c("06/29/07") ; d2 <- c("07/02/07")
D1 <- as.Date(d1, "%m/%d/%Y")
D2 <- as.Date(d2, "%m/%d/%Y")
D1 + 2 ; D1 - D2
difftime(D1,D2,units = "days")
```

1

3

5

或者改变显示形式:

```
> format(Sys.Date(), format="%A, %d %B %Y")
```

1

help(strptime)



导言

R 简介

- 初识 R
- 下载和安装
- 帮助系统
- 图形界面
- 包

数据处理

- 输入输出
- 数据结构
- 数据操作及变换
- 日期

字符处理

- 高级数据处理

函数及运算

- 数学函数
- 矩阵运算

编程

绘图

- 基本元素
- 图形设备

统计

- 分布与抽样
- 假设检验
- 拿“回归”来说事

有用的技巧

```
> paste("A", 1:6, sep = "")  
[1] "A1" "A2" "A3" "A4" "A5" "A6"  
> paste("A", 1:6, sep = '', collapse = '')  
[1] "A1A2A3A4A5A6"
```

1

3



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

substr(x, start, stop), strsplit(x, split)

```
> ?strsplit
> strReverse <- function(x)
+           sapply(lapply(strsplit(x, NULL), rev),
+                  paste, collapse="")
> strReverse(c("abc", "Statistics"))
[1] "cba"          "scitsitatS"
```

2

4

6



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符串处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

grep, grepl, sub, gsub, regexpr, gregexpr

```
> gsub("([ab])", "\\1_\\1_", "abc and ABC")  
[1] "a_a_b_b_c a_a_nd ABC"
```

2



apply(X, MARGIN, FUN, ...)

```
> d <- data.frame(x1 = rnorm(100), x2 = runif(100))
> apply(d, 2, quantile, seq(0, 1, by = 0.1))
```

	x1	x2
0%	-1.7786545	0.002695648
10%	-0.8978897	0.053191817
20%	-0.6261194	0.142789908
30%	-0.3419569	0.238808594
40%	-0.1210771	0.324463804
50%	0.1200293	0.432253790
60%	0.3624039	0.508909175
70%	0.7627071	0.772415243
80%	0.9944929	0.881434797
90%	1.5533980	0.949991713
100%	2.7399503	0.998079380

导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

10 函数及运算

数学函数

矩阵运算

12 编程

14 绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

by(data, INDICES, FUN), lapply, sapply, mapply

```
> tmp <- by(warpbreaks, tension,  
+           function(x) lm(breaks ~ wool, data = x))  
> sapply(tmp, coef)
```

	L	M	H
(Intercept)	44.55556	24.000000	24.555556
woolB	-16.33333	4.777778	-5.777778

2

4

6



tapply(X, INDEX, FUN = NULL), aggregate(x, by, FUN)

```
> head(warpbreaks,3)
  breaks wool tension
1      26     A      L
2      30     A      L
3      54     A      L
> tapply(warpbreaks$breaks,
+        list(warpbreaks$wool,warpbreaks$tension),mean)
           L           M           H
A 44.55556 24.00000 24.55556
B 28.22222 28.77778 18.77778
> aggregate(breaks ~ wool + tension,
+           data = warpbreaks, mean)
  wool tension  breaks
1    A      L 44.55556
2    B      L 28.22222
3    A      M 24.00000
4    B      M 28.77778
5    A      H 24.55556
6    B      H 18.77778
```

导言

R 简介

2

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

6

输入输出

数据结构

8

数据操作及变换

日期

字符处理

10

高级数据处理

函数及运算

12

数学函数

矩阵运算

14

编程

绘图

16

基本元素

图形设备

统计

18

分布与抽样

假设检验

拿“回归”来说事

有用的技巧



引言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

```
xtabs(formula = ., data = parent.frame(),  
subset)
```

```
> xtabs( breaks ~ wool + tension, data = warpbreaks)
      tension
wool    L    M    H
  A  401  216  221
  B  254  259  169

> tapply(warpbreaks$breaks,
+        list(warpbreaks$wool, warpbreaks$tension), sum)
      L    M    H
A  401  216  221
B  254  259  169
```

1

3

5

7

9



导言

R 简介

初识 R
下载和安装
帮助系统
图形界面
包

数据处理

输入输出
数据结构
数据操作及变换
日期
字符处理

高级数据处理

函数及运算

数学函数
矩阵运算

编程

绘图

基本元素
图形设备

统计

分布与抽样
假设检验
拿“回归”来说事

有用的技巧

```
> with(warpbreaks, table(wool, tension))
      tension
wool  L  M  H
   A  9  9  9
   B  9  9  9
```

`prop.table(),margin.table()`



导言

R 简介

- 初识 R
- 下载和安装
- 帮助系统
- 图形界面
- 包

数据处理

- 输入输出
- 数据结构
- 数据操作及变换
- 日期
- 字符处理

高级数据处理

函数及运算

- 数学函数
- 矩阵运算

编程

绘图

- 基本元素
- 图形设备

统计

- 分布与抽样
- 假设检验
- 拿“回归”来说事

有用的技巧

```
merge(x, y, by = intersect(names(x),
names(y)), by.x = by, by.y = by, all = FALSE,
all.x = all, all.y = all, sort = TRUE,
suffixes = c(".x", ".y"), incomparables = NULL,
...)
```



reshape

```
> head(df3)
  school class time      score
1      1     9    1 -0.0105537
2      1    10    1  1.3424263
3      1     9    2 -0.3533924
4      1    10    2 -0.9512261
5      2     9    1 -1.1927487
6      2    10    1 -0.5784864

> reshape(df3, idvar=c("school", "class"),
+         direction="wide")
  school class      score.1      score.2
1      1     9 -0.0105537 -0.3533924
2      1    10  1.3424263 -0.9512261
5      2     9 -1.1927487 -0.4581521
6      2    10 -0.5784864 -1.0176699
9      3     9  1.6342330 -0.5758275
10     3    10  0.6874732  1.2867034
```

导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

- ① `+, -, ×, ÷, ^, %%, %/%`
- ② `< > <= >= == .. != ..`
- ③ `sin, cos, tan, asin, acos, atan, log, log10, exp`
- ④ `max(x), min(x)`
- ⑤ `range(x)`
- ⑥ `sum(x)`
- ⑦ `diff(x)`
- ⑧ `prod(x)`
- ⑨ `mean(x)`
- ⑩ `abs(x)`
- ⑪ `sqrt(x)` $x^{0.5}$
- ⑫ `median(x)`

[导言](#)[R 简介](#)[初识 R](#)[下载和安装](#)[帮助系统](#)[图形界面](#)[包](#)[数据处理](#)[输入输出](#)[数据结构](#)[数据操作及变换](#)[日期](#)[字符处理](#)[高级数据处理](#)[函数及运算](#)[数学函数](#)[矩阵运算](#)[编程](#)[绘图](#)[基本元素](#)[图形设备](#)[统计](#)[分布与抽样](#)[假设检验](#)[拿“回归”来说事](#)[有用的技巧](#)



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

- ① `quantile(x, probs=)`
- ② `IQR(x)`
- ③ `weighted.mean(x, w)`
- ④ `rank(x)`
- ⑤ `var(x)`
- ⑥ `cor(x, y)`
- ⑦ `round(x, n)`
- ⑧ `log(x, base)`
- ⑨ `scale(x)`
- ⑩ `integrate(f, lower, upper)`
- ⑪ `pmin(x, y, ...)` , `pmax(x, y, ...)`



导言

R 简介

- 初识 R
- 下载和安装
- 帮助系统
- 图形界面
- 包

数据处理

- 输入输出
- 数据结构
- 数据操作及变换
- 日期
- 字符处理
- 高级数据处理

函数及运算

数学函数

- 矩阵运算

编程

绘图

- 基本元素
- 图形设备

统计

- 分布与抽样
- 假设检验
- 拿“回归”来说事

有用的技巧

- 1 `cumsum(x)`
- 2 `cumprod(x)`
- 3 `union(x,y)` $x \cup y - x \cap y$
- 4 `intersect(x,y)` $x \cap y$
- 5 `setdiff(x,y)` $x - x \cap y$
- 6 `filter(x, filter)`
- 7 `fft(x)`
- 8 `mvfft(x)`
- 9 `convolve(x,y)`
- 10 ...



导言

R 简介

初识 R
下载和安装
帮助系统
图形界面
包

数据处理

输入输出
数据结构
数据操作及变换
日期
字符处理
高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素
图形设备

统计

分布与抽样
假设检验
拿“回归”来说事

有用的技巧

- **t(x)**
- **diag(x)** 对角阵
- **%*%** 矩阵运算
- **solve(a,b)** $a \%*\% x = b$
- **solve(a)**
- **eigen(x)**
- **rowSums(x), colsum(x)**
- **rowMeans(x), colMeans(x)**
- **dist(x)**



导言

R 简介

- 初识 R
- 下载和安装
- 帮助系统
- 图形界面
- 包

数据处理

- 输入输出
- 数据结构
- 数据操作及变换
- 日期
- 字符处理
- 高级数据处理

函数及运算

- 数学函数
- 矩阵运算

编程

绘图

- 基本元素
- 图形设备

统计

- 分布与抽样
- 假设检验
- 拿“回归”来说事

有用的技巧

```
if (test)
{
    ...true statements...
}
else
{
    ...false statements...
}
```

```
for(var in seq) expr,while(cond) expr,repeat
expr,break,next,ifelse
```

1

3

5

7



导言

R 简介

初识 R
下载和安装
帮助系统
图形界面
包

数据处理

输入输出
数据结构
数据操作及变换
日期
字符处理
高级数据处理

函数及运算

数学函数
矩阵运算

编程

绘图

基本元素
图形设备

统计

分布与抽样
假设检验
拿“回归”来说事

有用的技巧

```
gg <- function(n)
{
  a <- numeric(n)
  a[1] <- 1
  a[2] <- 1
  for(i in 3:n){
    a[i] <- a[i-1] + 2*a[i-2] + 1
  }
  return(a)
}
> gg(9)
[1] 1 1 4 7 16 31 64 127 256
```




导言

R 简介

初识 R
下载和安装
帮助系统
图形界面
包

数据处理

输入输出
数据结构
数据操作及变换
日期
字符处理
高级数据处理

函数及运算

数学函数
矩阵运算

编程

绘图

基本元素
图形设备

统计

分布与抽样
假设检验
拿“回归”来说事

有用的技巧

“高水平” (High_level)

图形设备上绘制新图

“低水平 (Low_level) ”

已经存在图形上添加更多的绘图信息，如点、线、多边形等

“交互式” (Interactive)

可以使用如鼠标这类的定点装置来添加或提取绘图信息



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

	名称	描述
屏幕 显示	x11 windows	X 窗口 Windows 窗口
文件 设备	postscript pdf pictex png jpeg bmp xfig win.metafile	ps 格式文件 pdf 格式文件 供 L ^A T _E X 使用的文件 png 格式文件 jpeg 格式文件 bmp 格式文件 供 XFIG 使用的图形格式 emf 格式的文件



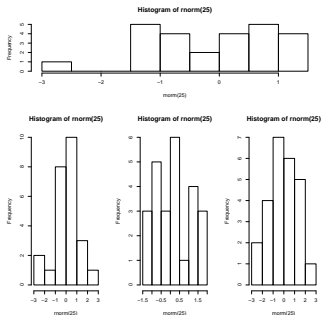
```
nf <- layout(matrix(c(1, 1, 1,
                      2, 3, 4,
                      2, 3, 4),nr = 3, byrow = T))

hist(rnorm(25));      hist(rnorm(25));
hist(rnorm(25));      hist(rnorm(25))
```

2

[导言](#)
[R 简介](#)

4

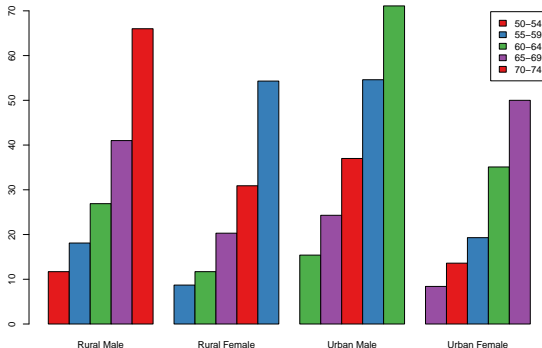
[初识 R](#)
[下载和安装](#)
[帮助系统](#)
[图形界面](#)
[包](#)
[数据处理](#)
[输入输出](#)
[数据结构](#)
[数据操作及变换](#)
[日期](#)
[字符处理](#)
[高级数据处理](#)
[函数及运算](#)
[数学函数](#)
[矩阵运算](#)
[编程](#)
[绘图](#)
[基本元素](#)
[图形设备](#)
[统计](#)
[分布与抽样](#)
[假设检验](#)
[拿“回归”来说事](#)
[有用的技巧](#)




```
> library(RColorBrewer)
> v.col <- brewer.pal(4,"Set1")
> barplot(VADeaths, beside = TRUE,col = v.col,
          legend.text=row.names(VADeaths))
```

1

3



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧



导言

R 简介

- 初识 R
- 下载和安装
- 帮助系统
- 图形界面
- 包

数据处理

- 输入输出
- 数据结构
- 数据操作及变换
- 日期
- 字符处理
- 高级数据处理

函数及运算

- 数学函数
- 矩阵运算

编程

绘图

- 基本元素
- 图形设备

统计

分布与抽样

- 假设检验
- 拿“回归”来说事

有用的技巧

sample(n)	随机组合 $1, \dots, n$
sample(x)	随机组合向量 x , $length(x) > 1$
sample(x, replace = T)	解靴带法
sample(x, n)	非放回的从 x 中抽取 n 项
sample(x, n, replace = T)	放回的从 x 中抽取 n 项
sample(x, n, replace = T, prob = p)	以概率 p , 放回的从 x 中抽取 n 项



导言

R 简介

初识 R
下载和安装
帮助系统
图形界面
包

数据处理

输入输出
数据结构
数据操作及变换
日期
字符处理
高级数据处理

函数及运算

数学函数
矩阵运算

编程

绘图

基本元素
图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

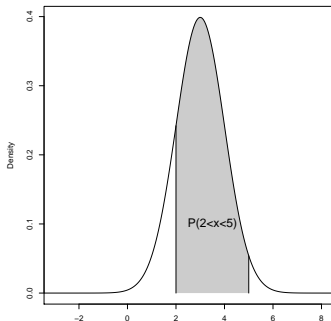
有用的技巧

分布相关

如已知 $X \sim N(3, 1)$, 计算 $P(2 \leq X \leq 5)$

```
> pnorm(5, 3, 1) - pnorm(2, 3, 1)
[1] 0.8185946
```

2



d: 概率密度函数 | q: 分位数函数

p: 累积分布函数 | r: 该分布的随机数



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

分布	R 函数	附加参数	默认参数
beta	beta	$\text{shape1}(\alpha), \text{shape2}(\beta)$	
二项	binom	$\text{size}(n), \text{prob}(p)$	
χ^2	chisq	df	
均匀	unif	$\text{min}(a), \text{max}(b)$	$\text{min} = 0, \text{max} = 1$
指数	exp	rate	$\text{rate} = 1$
F	f	$\text{df1}(r_1), \text{df2}(r_2)$	
伽玛	gamma	$\text{shape}(\alpha), \text{scale}(\theta)$	$\text{scale} = 1$
超几何	hyper	$m = N_1, n = N_2, k = n$	
正态	norm	$\text{mean}(\mu), \text{sd}(\sigma)$	$\text{mean} = 0, \text{sd} = 1$
泊松	pois	$\text{lamda}(\lambda)$	
t	t	df	
威布尔	weibull	$\text{shape}(\alpha), \text{scale}(\theta)$	$\text{scale} = 1$



导言

R 简介

初识 R
下载和安装
帮助系统
图形界面
包

数据处理

输入输出
数据结构
数据操作及变换
日期
字符处理
高级数据处理

函数及运算

数学函数
矩阵运算

编程

绘图

基本元素
图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

```
> apropos(".test")
[1] ".valueClassTest"      "ansari.test"
[3] "bartlett.test"         "binom.test"
[5] "Box.test"              "chisq.test"
[7] "cor.test"              "file_test"
[9] "fisher.test"           "fligner.test"
[11] "friedman.test"         "kruskal.test"
[13] "ks.test"               "mantelhaen.test"
[15] "mauchley.test"         "mauchly.test"
[17] "mcnemar.test"          "mood.test"
[19] "oneway.test"           "pairwise.prop.test"
[21] "pairwise.t.test"       "pairwise.wilcox.test"
[23] "poisson.test"          "power.anova.test"
[25] "power.prop.test"       "power.t.test"
[27] "PP.test"               "prop.test"
[29] "prop.trend.test"       "quade.test"
[31] "shapiro.test"          "t.test"
[33] "var.test"              "wilcox.test"
```




导言

R 简介

- 初识 R
- 下载和安装
- 帮助系统
- 图形界面
- 包

数据处理

- 输入输出
- 数据结构
- 数据操作及变换
- 日期
- 字符处理
- 高级数据处理

函数及运算

- 数学函数
- 矩阵运算

编程

绘图

- 基本元素
- 图形设备

统计

- 分布与抽样
- 假设检验

拿“回归”来说事

有用的技巧

拿常见的 `lm`, `glm` 模型来说, `y ~ model` 是一种特定的格式, 表示以 `y` 为响应变量, 模型为 `model`。其中 `model` 中的变量由 `+` 来连接, 或者由 `:` 来表示变量间的“交互作用”。除了 `+` 和 `:`, 我们使用 `*` 来表示 '`a + b + a:b`'。 `(a + b + c)^2` 表示 `(a + b + c) * (a + b + c)`, 即主因素 `a`、`b`、`c` 和各个因素的交互作用。`-` 表示去掉之意。 `(a + b + c)^2 - a:b` 表示 '`a + b + c + b:c + a:c`'。在公式表达中除了变量和因子名外, 运算符号也是可以存在的。如 '`log(y) a + log(x)`' 是合法的。

符号. 在 `update` 函数中有特殊的意义, 它表示“已经存在”之意。

```
fm <- aov(Speed ~ Run + Expt)
fm0 <- update(fm, . ~ . - Run)
```

2

Example



导言

R 简介

初识 R
下载和安装
帮助系统
图形界面
包

数据处理

输入输出
数据结构
数据操作及变换
日期
字符处理
高级数据处理

函数及运算

数学函数
矩阵运算

编程

绘图

基本元素
图形设备

统计

分布与抽样
假设检验

拿“回归”来说事

有用的技巧

```
> x <- 1:20
> y <- x + rnorm(20,0,0.1)
> lm.y <- lm(y ~ x)
> lm.y
Call:
lm(formula = y ~ x)

Coefficients:
(Intercept)          x
   -0.06118      1.00399

> names(lm.y)
[1] "coefficients"  "residuals"    "effects"      "rank"
[5] "fitted.values" "assign"        "qr"           "df.residual"
[9] "xlevels"       "call"         "terms"        "model"
```



导言

R 简介

初识 R
下载和安装
帮助系统
图形界面
包

数据处理

输入输出
数据结构
数据操作及变换
日期
字符处理
高级数据处理

函数及运算

数学函数
矩阵运算

编程

绘图

基本元素
图形设备

统计

分布与抽样
假设检验

拿“回归”来说事

有用的技巧

Function	Meaning
summary()	summary of the fitted model
coef()	estimated model parameters
resid()	model residuals of the fitted model
fitted()	the fitted values of the model
deviance()	residual sum of squares
anova()	anova table
predict()	predictions
plot()	diagnostic plot



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

在 2.5.0 版本中，R 开始命令自动补全功能：

- 单击 **Tab**，自动补全 R 命令；
- 两次 **Tab**，返回所有可能的补全命令列表。



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

- **Hmisc** 包中的 **latex()** 函数
- **xtable** 包中的 **xtable()** 函数
- **quantreg** 包中的 **latex.table()**

```
a <- matrix(1:6, nr=1) # require(xtable)
colnames(a) <- paste("col", 1:6)
xtable(a)
```

2

给自己的循环加个 Progress



导言

R 简介

初识 R
下载和安装
帮助系统
图形界面
包

数据处理

输入输出
数据结构
数据操作及变换
日期
字符处理
高级数据处理

函数及运算

数学函数
矩阵运算

编程

绘图

基本元素
图形设备

统计

分布与抽样
假设检验
拿“回归”来说事

有用的技巧

```
testit <- function(x = sort(runif(20)), ...)  
{  
  pb <- txtProgressBar(...)  
  for(i in c(0, x, 1)) {Sys.sleep(0.5);  
    setTxtProgressBar(pb, i)}  
  Sys.sleep(1)  
  close(pb)  
}  
testit()
```

- eclipse
- Tinn-R
- RWinEdt
- Emacs
- UltraEdit



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

Q&A



导言

R 简介

- 初识 R
- 下载和安装
- 帮助系统
- 图形界面
- 包

数据处理

- 输入输出
- 数据结构
- 数据操作及变换
- 日期
- 字符处理
- 高级数据处理

函数及运算

- 数学函数
- 矩阵运算

编程

绘图

- 基本元素
- 图形设备

统计

- 分布与抽样
- 假设检验
- 拿“回归”来说事

有用的技巧



导言

R 简介

初识 R

下载和安装

帮助系统

图形界面

包

数据处理

输入输出

数据结构

数据操作及变换

日期

字符处理

高级数据处理

函数及运算

数学函数

矩阵运算

编程

绘图

基本元素

图形设备

统计

分布与抽样

假设检验

拿“回归”来说事

有用的技巧

Q&A

- Email: [sunbjt<at>gmail.com](mailto:sunbjt@gmail.com)
- Blog : <http://www.bjt.name>

Jump to first slide