Data Analysis Using R: Chapter07

罗智超 (ROKIA.ORG)

1 通过本章你将学会

- 掌握数学函数
- 掌握统计函数
- 掌握字符函数
- 使用循环
- 条件语句
- 编写自己的函数
- 错误捕获
- 调试函数

2 Mathematal Functions

– abs(x)

– sqrt(x)

– ceiling(x)

– floor(x)

– trunc(x)

– round(x,digits=n)

– signif(x,digits=n)

– cos(x),sin(x),tan(x)

– acos(x),asin(x),atan(x)

– cosh(x), sinh(x), tanh(x)

– acosh(x), asinh(x), atanh(x)

– log(x,base=n) log(x) log10(x) exp(x)

3

- ceiling takes a single numeric argument x and returns a numeric vector containing the smallest integers not less than the corresponding elements of x.
- floor takes a single numeric argument x and returns a numeric vector containing the largest integers not greater than the corresponding elements of x.
- trunc takes a single numeric argument x and returns a numeric vector containing the integers formed by truncating the values in x toward 0.
- round rounds the values in its first argument to the specified number of decimal places (default 0).
- signif rounds the values in its first argument to the specified number of significant digits.

4 Statistical Functions

– mean(x, trim = 0.1, na.rm=TRUE)

– median(x)

– sd(x)

- var(x)
- mad(x) Median Absolute Deviation
- quantile(x, probs)
- range(x)
- sum(x)
- diff(x,lag=1)
- min(x) max(x)
- scale(x, center=TRUE, scale=TRUE)

5 Probability Functions

5.1 dpqr

- dpqr:distribution_abbreviation
- d=density
- p=distribution function
- q=quantile function
- r=random generation (random deviates)

6

6.1 Probability distributions

- The common probability functions are listed below

Dist	Abbre	Dist	Abbre
Beta	beta	Logistic	logis

Dist	Abbre	Dist	Abbre
Binomial	binom	Multinomial	multinom
Cauchy	cauchy	Negative binomial	nbinom
Chi-quared (noncentral)	chisq	Normal	norm
Exponential	exp	Poisson	pois

7

Dist	Abbre	Dist	Abbre
F	f	Wilcoxon Signed Rank	signrank
Gamma	gamma	T	t
Geometric	geom	Uniform	unif
Hypergeometric	hyper	Weibull	weibull
Lognormal	lnorm	Wilcoxon Rank Sum	wilcox

8

```
#Plot the standard normal curve  on the interval  [-3,3]
x <- pretty(c(-3,3), 30)
y <- dnorm(x)
plot(x, y, type='l',xlab="Normal Deviate",
     ylab="Density",yaxs = "i" )

#What is the area under the standard normal
#curve to the right of z = 1.96?
pnorm(1.96)

#What is the value of the 90th percentile
#of a normal distribution with
#a mean of 500 and a standard deviation of 100?
```

```
qnorm(.9, mean=500, sd=100)

#Generate 50 random normal deviates with
#a mean of 50 and a standard deviation of 10.

rnorm(50, mean=50, sd=10)
```

9

9.1 SETTING THE SEED FOR RANDOM NUMBER GENERATION

```
runif(5)
runif(5)

set.seed(1234)
runif(5)

set.seed(1234)
runif(5)
```

10

10.1 GENERATING MULTIVARIATE NORMAL DATA

- The mvrnorm function in the MASS package
- mvrnorm(n, mean, sigma)
- n is the desired sample size

- mean is the vector of means
- sigma is the variance-covariance (or correlation) matrix

11

```
library(MASS)
mean <- c(230.7, 146.7, 3.6)
sigma <- matrix(c(15360.8, 6721.2, -47.1, 6721.2,
                  4700.9,-16.5, -47.1, -16.5,
                  0.3),nrow=3, ncol=3)
set.seed(1234)
mydata <- mvrnorm(500, mean, sigma)
mydata <- as.data.frame(mydata)
names(mydata) <- c("y","x1","x2")
head(mydata)
```

12

12.1 Character functions

- nchar(x)
- substr(x, start, stop)
- grep(pattern, x, ignore.case=FALSE,fixed=FALSE)
- sub(pattern, replacement, x, ignore.case=FALSE, fixed=FALSE)
- strsplit(x, split)
- paste(..., sep="")
- toupper(x)
- tolower(x)

13

13.1 Other useful functions

length(x) seq(from , to, by) rep(x, ntimes) cut(x, n) pretty(x, n) cat(??)

14

14.1 Control flow

14.1.1 Repetition and looping

- for (var in seq) statement
- while (cond) statement

14.1.2 Conditional execution

- if (cond) statement
- if (cond) statement1 else statement2
- ifelse(cond, statement1, statement2)
- switch(expr, …)

15 User-written functions

```
myfunction <- function(arg1, arg2, ... )
    { statements return(object) }
```

16 Google's R Style Guide

- Google's R Style Guide

17 操作环境对象

函数	描述
assign(x,value)	在 envir 环境中将名称 x 赋给 value 对象
get(x)	在 envir 环境中获得与名称关联的对象
exists()	判断在环境 envir 中是否定义了名称 x
attach()	将列表、数据框或者数据文件中的对象添加到当前的搜索路径
detach()	将列表、数据框或者数据文件中的对象从当前的搜索路径中删除

18 调试函数

- tryCatch()