

## 《易学统计》常用R语言表达式、函数、操作符范例

操作符:

数学运算操作符: + - \* /

矩阵、数组: (1:10) c(1,2,3,4,5) x[1:3] x[a>0] x[r>10, w>4]

逻辑操作符: > < >= <= == != & | is.na(X) !is.na(X)

函数:

- abs(x): absolute value
- log(x): natural logarithm
- log10(x): common logarithm
- log2(x): 2 based logarithm
- exp(x):  $e^x$
- sqrt(x): square root
- ceiling(x): ceiling(3.475)=4
- floor(x): floor(3.475)=3
- trunc(x): trunc(5.99)=5
- round(x, digits=n): round(3.475, digits=2)=3.48
- signif(x, digits=n): signif(3.475, digits=2)=3.5
- cos(x)、sin(x)、tan(x)、acos(x)、cosh(x)、acosh(x), etc.

条件表达式: X[A>60 & !is.na(A)] <- (X-10)[A>60 & !is.na(A)]

输入公式例:

1、从日期变量 x="5/12/2009" 中求出月份、日期、年份:

- a) 月份: unlist(strsplit(x, "/"))[1]
- b) 日期: unlist(strsplit(x, "/"))[2]
- c) 年份: unlist(strsplit(x, "/"))[3]

2、按记录顺序产生记录系列号: (1:nrow(EmpowerStatsR))

3、产生随机数:

可以重复的从 1 到 100 之间的随机整数:

```
sample(1:100, nrow(EmpowerStatsR), replace=T)
```

不重复的随机整数:

```
sample(1:nrow(EmpowerStatsR), nrow(EmpowerStatsR), replace=F)
```

0 到 1 之间的随机数: runif(nrow(EmpowerStatsR))

4、变量按比例重新赋值, 如按 0-1 的取值范围重新取值

```
x.new <- scale(x,  
               center=min(x, na.rm=TRUE),  
               scale =diff(range(x, na.rm=TRUE)))
```

5、计算每个值对应的百分位数

```
x.new <- scale(rank(x,  
  ties.method="average",  
  na.last="keep"),  
  center=1,  
  scale=sum(!is.na(x))-1)
```

6、合并两个或多个变量字符，如：

```
paste("abc", 123, sep="") = "abc123"  
paste(var1, var2, ..., varN, sep="")
```