

Тема 8.

**R - язык и вычислительная среда
для обработки информации
современными статистическими
методами**

Примеры использования

<http://mpoctok.narod.ru/r/intro.htm>

Введение

В R используется интерфейс командной строки

В 2010 году R вошёл в список победителей конкурса журнала InfoWorld в номинации на лучшее открытое программное обеспечение для разработки приложений

Особенности

R поддерживает широкий спектр статистических и численных методов и обладает хорошей расширяемостью с помощью пакетов.

Пакеты представляют собой библиотеки для работы специфических функций или специальных областей применения.

Важной особенностью R являются графические возможности, заключающиеся в возможности создания качественной графики, которая может включать математические символы.

Запуск R

Способы работы в системе R:

"с терминала": запускается программой Rterm, при этом вы работаете как в DOS вводя команды в ответ на запрос системы; вы можете загрузить с диска и файлы, содержащие команды языка R и описания функций-процедур. Выход из этого режима командой quit() или просто q().

"используя графический интерфейс": запускается программой Rgui, при этом открывается окно со списком различных меню, пиктограммами и окном ввода-вывода, в котором можно работать как с терминала.

Просмотр списка объектов

Войдя в систему **R** можно посмотреть что у вас есть в рабочей области.

Для этого введите команду **ls()**, по которой на экран будет выведен список всех ОБЪЕКТОВ, находящихся в рабочей области: *простых переменных, структур, наборов данных, процедур.*

Если вы хотите удалить объект под именем xxx, то введите **rm(xxx)**.

Справка

Справку о любой команде вы получите набрав **?cmd** или **help(cmd)** или **help('cmd')**.

Например:

?ls или

help(ls) или

help("ls") или

help('ls').

На панели R вверху экрана есть клавиша Help

Чтение из файла

КАК ПРОЧИТАТЬ ИЗ ФАЙЛА ДАННЫЕ,
РАЗДЕЛЁННЫЕ ЗАПЯТОЙ

```
dat<- read.table(file="f.txt", sep=",")
```

```
# создаёт объект-таблицу dat
```

```
# в строках которой по записям размещены данные
```

```
# из файла f.txt
```

```
# которые должны быть разделены запятой
```

Вычисление среднего

```
# В переменную a поместить список всех оценок:
```

```
a <- c(4,3,3,3,3,4,4,4,4,4,5,4,4,4,5,5,5,5,+  
3,5,5,4,4,3,3,4,4,3,5,5,4,3,3,4,4,3,3,5,4,5,5)
```

```
# В переменную n поместить количество оценок:
```

```
length(a) -> n
```

```
# Средний балл:
```

```
m <- mean(a)
```

Преобразование в таблицу

Таблица (горизонтальная; вариационный ряд)
с подсчётом количества оценок:

```
t <- table(a)
```

Преобразование в более удобный формат данных
(вертикальную таблицу):

```
f <- as.data.frame(t)
```

(t(a) – обычное транспонирование)

Вычисление процентной доли
и запись её в третий столбец:

```
mapply(function(r) r*100/n, f[,2]) -> f[,3]
```

Вывод результатов и результаты

```
# Заголовки столбцов:  
colnames(f) <- c("Оценка", "Кол-во", "%")
```

```
# Вывод результатов:
```

```
a
```

```
n
```

```
m
```

```
f
```

```
[1] 4 3 3 3 3 4 4 4 4 4 5 4 4 4 5 5 5 5 3 5 5 4 4 3 3 4 4 3 5 5 4 3 3 4 4 3 3 5
```

```
[39] 4 5 5
```

```
[1] 41
```

```
[1] 4
```

```
Оценка Кол-во %
```

```
1 3 12 29.26829
```

```
2 4 17 41.46341111
```

```
3 5 12 29.26829
```

Некоторые операторы

Вычисление частотей (добавление 3-го столбца)

```
t[,3] <- 100*t/sum(t) ?
```

Вычисление статистических характеристик:

```
summary (t(a))
```

КАК НАРИСОВАТЬ ГИСТОГРАММУ ЗНАЧЕНИЙ ЭЛЕМЕНТОВ ВЕКТОРА X :

```
hist(x)           # нарисовали гистограмму
```

```
hist(x,breaks=11)  
# нарисовали гистограмму для 10 ячеек
```

Графика

КАК НАРИСОВАТЬ ПРОИЗВОЛЬНЫЙ ГРАФИК:

```
plot (x,y,type="l")  или  
...      ="p"
```

ПОЛИГОН ЧАСТОТ:

```
plot (t,type="l",main="Полигон частот",xlab="x",ylab="y")
```

КАК ВВЕСТИ ДАННЫЕ С КЛАВИАТУРЫ:

```
x<- readline("Введите номер. ")
```

```
# x равно символьной строке, которую Вы введёте  
# после надписи на экране "Введите номер."
```

Условие, цикл

`if(cond) expr`

`while(cond) expr`

`for(var in seq) expr`
`for(i in 1:5)`