

Biometry. Lecture 7

Alexey Shipunov

Minot State University

February 1, 2012

Outline

- 1 Questions and answers
- 2 Processing data
 - Entering data into R
 - Overgrown calculator
 - The basics of R graphics

Outline

- 1 Questions and answers
- 2 Processing data
 - Entering data into R
 - Overgrown calculator
 - The basics of R graphics

Previous final question: the answer

File `eq.txt` has tabs instead of semicolons as delimiters.
How to use `read.table()` command in this case?

```
> read.table("data/eq.txt") # or  
> read.table("data/eq.txt", sep="\t")
```

Both ways are equal if your data has no spaces in it. If spaces are present, the second way is preferable.

1st exam

- Paper exam
- 20 questions/tasks: multiple choice and short answers
- You may use lab materials and computers

Processing data

Entering data into R

How to start and to finish

```
> getwd()
> setwd(...)
(Or do it through menu: File > Change dir)
> savehistory("<name>.r")
```

Saving workspace

```
> save.image("20110127.rd")  
> load("20110127.rd")
```

If you save the image, you may the restore all object created during the R session. The image will be saved in R binary format.

What `q()` command is doing

If you answer “yes” on the last question, two files will be created:

- `RData` will contain all your objects
- `Rhistory` will contain command history
- If you are bored with this question, you may want to enter `q("no")`

When you start R from the same folder, these two files will be automatically loaded and your objects and history will be automatically restored.

Processing data

Overgrown calculator

Vectorization and brackets

```
> log(((sqrt(sum(c(2,2))))^2)*2.5)
```

Calculation order: from inside to outside

The order of calculations

```
> 2+3*5  
but  
> 2+(3*5)  
is better!
```

Distributions

```
> rnorm(15)
```

`rnorm()` is one of almost 20 embedded distributions. By default, mean is 0 and sd is 1. You may change it, e.g., `rnorm(15, mean=10)` will give numbers around 10. Normal distribution is a result of the influence of *multiple independent random factors*, dart game is a good example.

Processing data

The basics of R graphics

Simple plot

```
> plot(1:20)
```

Title and legend

```
> plot(1:20, main="My very important title")  
> legend("topleft", pch=1, legend="My precious dots")
```

`legend()` needs to “understand” what (color, points etc.) to describe; `pch` is a type of points. You may use `plot(1:20, pch=2)` to have triangles as dots.

Two types of graphical commands: plotting and adding

```
> plot(cars)  
> title(main="Cars of 1920s")
```

`cars` is an embedded data, run `?cars` for explanation

plot () is a smart (generic) command

```
> plot(cars)
> plot(trees)
> plot(uspop)
> plot(HairEyeColor)
```

Types of `plot()`

```
> plot(uspop, type="p")  
> plot(uspop, type="l")  
> plot(uspop, type="c")  
> plot(uspop, type="s")  
> plot(uspop, type="o")  
> plot(uspop, type="b")
```

Empty plot with added points and grid

```
> plot(1:20, type="n")  
> points(1:20, 1:20, pch=2, col=2)  
> grid(5,5)
```

Empty `plot()` will make a coordinate grid. This is frequently used if you want to construct a complex graphs.

Graphical devices

```
> plot(1:20)  
> dev.off()
```

`dev.off()` will close the current device

PDF graphical device

```
> pdf(file="1.pdf")  
> plot(1:20)  
> dev.off()
```

PDF format is appropriate for the inclusion in reports, especially if you need to scale images

PNG graphical device

```
> png(file="1.png")  
> plot(1:20)  
> dev.off()
```

PNG format is more appropriate for the Web pages, it will not scale well

How to save current graph into the file

```
> plot(1:20)
> dev.copy(png, "2.png")
> dev.off()
```

The file will not be written on disk until you run `dev.off()`. On Windows, you may use a menu from graphical window.

Graphical options

```
> oldpar <- par(mfrow=c(2,1))  
> hist(cars$speed)  
> hist(cars$dist)  
> par(oldpar)
```

`mfrow` by default is `c(1,1)`

`par()` should be kept in the object and then restored

Interactive graphics

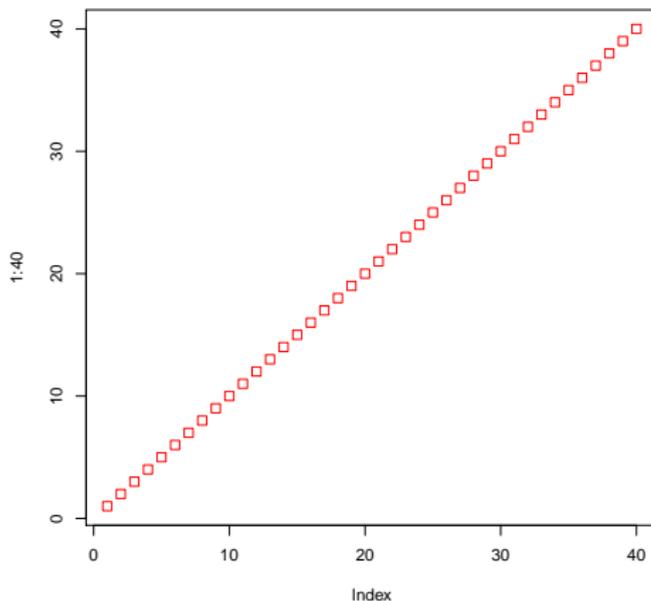
```
> plot(1:20)  
> text(locator(), "My beloved point", pos=4)
```

Click left mouse button, then right mouse button

Final question (3 points)

Final question (3 points)

Which command will produce this plot?



Summary: most important commands

- `plot()` draws plots
- `par()` regulates plots parameters

For Further Reading



A. Shipunov.

Biometry [Electronic resource].

2012—onwards.

Mode of access: `http:`

`//ashipunov.info/shipunov/school/biol_299`



P. Dalgaard

Introductory Statistics with R. 2nd edition.

Springer, 2008.

Section 1.2.