Biometry. Lecture 6

Alexey Shipunov

Minot State University

February 12, 2014

Outline

1. Questions and answers
2. The basics of R graphics
   - plot() command
3. Types of data
   - Measurement data

Outline

1. Questions and answers
2. The basics of R graphics
   - plot() command
3. Types of data
   - Measurement data

Outline

1. Questions and answers
2. The basics of R graphics
   - plot() command
3. Types of data
   - Measurement data

Exam 1

Starting...

```
> setwd("<working folder>")
```

or

“Change dir”

in menu!

Commands to look around

```
> dir() # shows files in working folder
> file.show() # shows content of file
> ls() # lists all objects
> str() # shows the structure of object
> head() # shows first rows of table object (data frame)
```

The basics of R graphics

plot() command

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 dist 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

Types of data

Measurement data

Measurement data

- For any two measurements, the third between them also has sense
- Best example: location on the ruler. Continuous, could be zero, positive and negative.
- Temperature has a restriction: there is a minimal temperature
- Angle is worse: there are both minimal and maximal angles

Discrete measurement data: counts

- Tis is the other kind of measurement data
- Number of items is always a whole number so there is the third between 2 and 4
- But the third number between 2 and 3 is a nonsense

“Parametric” and “non-parametric” data

- (a) Only *continuous measurement* data may be parametric
- In addition, parametric methods require: (b) suspected *normal distribution* of data and (c) sample ≥ 30
- Everything else should be studied with non-parametric methods

Measurement data in R

```
> x <- c(174, 162, 188, 192, 165, 168, 172)
> str(x)
 num [1:7] 174 162 188 192 165 168 172
> is.numeric(x)
 [1] TRUE
> is.vector(x)
 [1] TRUE
```

Some rules about vectors

- For every type of R object, there are functions `is.<something>()` and `as.<something>()` (e.g., `as.vector()` and `as.numeric()` will convert to vector and to numeric vector, respectively).
- Object names must not start with a number
- R is case-sensitive
- Please avoid to use names of popular functions (like `c()`) and keywords: `T` (TRUE), `F` (FALSE), `NA` (missing data), `NaN` (not a number), `Inf` (result of dividing by zero), `pi`

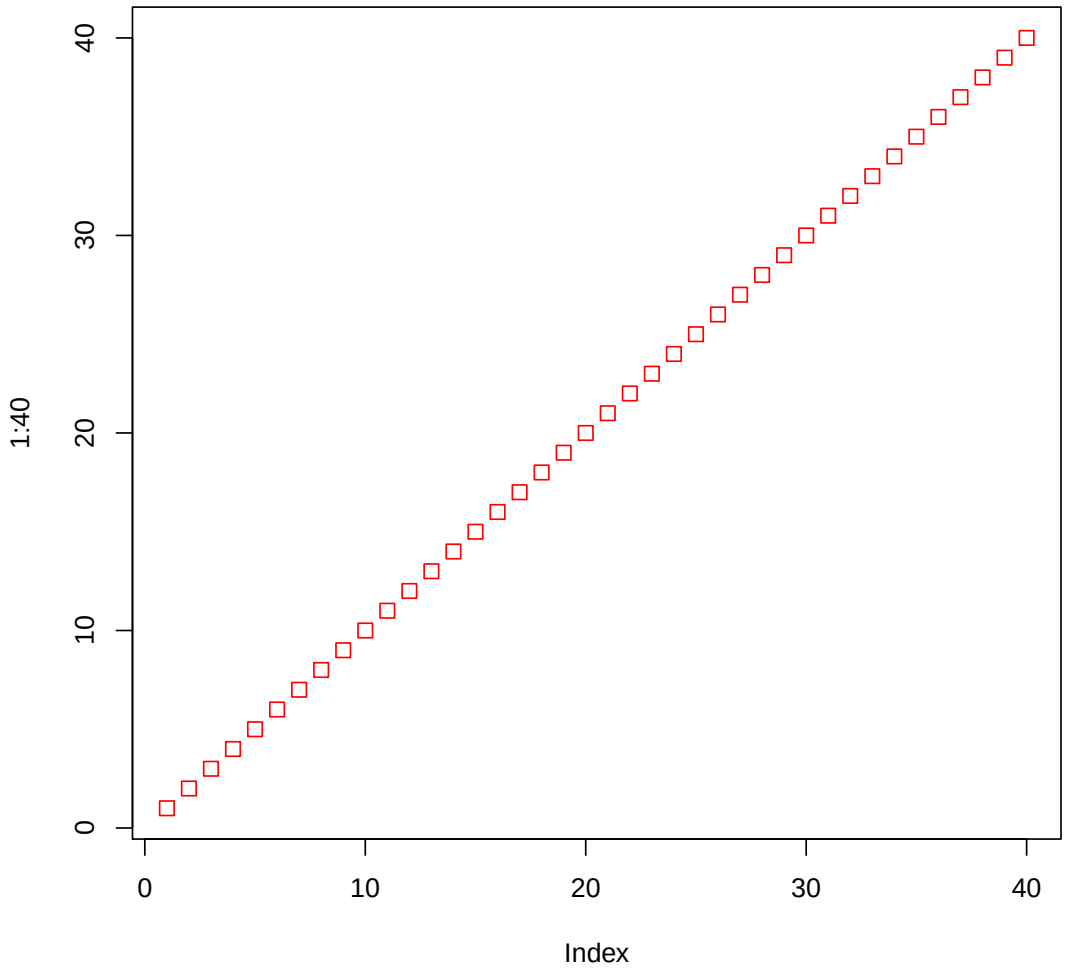
Finishing...

```
>savehistory("20140212.r")
```

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
- `str()`—shows the structure of object

For Further Reading

A. Shipunov.
Biometry [Electronic resource].
2012—onwards.
Mode of access:
http://ashipunov.info/shipunov/school/biol_240

A. Shipunov, and others.
Visual statistics. Use R!
DMK Press, 2012. [Translating from Russian.]