

# Biometry. Lecture 7

Alexey Shipunov

Minot State University

February 11, 2015



- 1 Questions and answers
- 2 Basics of R
  - The basics of R graphics
- 3 Types of data
  - Measurement data
- 4 Types of data
  - Ranked data



- 1 Questions and answers
- 2 Basics of R
  - The basics of R graphics
- 3 Types of data
  - Measurement data
- 4 Types of data
  - Ranked data



- 1 Questions and answers
- 2 Basics of R
  - The basics of R graphics
- 3 Types of data
  - Measurement data
- 4 Types of data
  - Ranked data



- 1 Questions and answers
- 2 Basics of R
  - The basics of R graphics
- 3 Types of data
  - Measurement data
- 4 Types of data
  - Ranked data



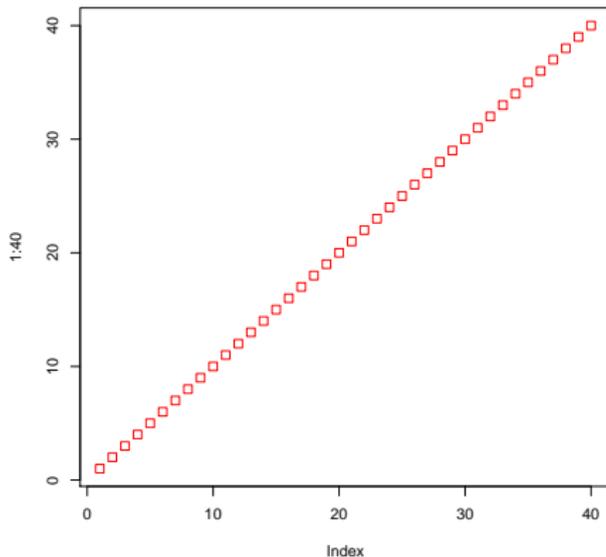
```
> setwd("<working folder>")  
or  
"Change dir"  
in menu!
```

On Mac, be sure that startup option is working: `getwd()`  
(`getwd()` checks if R is in working folder, `dir()` checks the folder  
content)



# Previous last question

Which command will produce this plot?



```
plot(1:40, pch=0, col=2)
```



# Enhance the plot from Lab 3

```
> scatter.smooth() # plots the line  
> log10() # logarithmic transformation
```



# Basics of R

## The basics of R graphics



# How to save current plot 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



# 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

- This 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  $\geq 30$
- 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
```



# Types of data

## Ranked data



# What if we cannot measure?

- In this case, we can use scale-like representation
- E.g., we can rank the student success from 1 to 5 (“very bad” to “excellent”)
- Or softness of mattress from 0 to 10 (“hard as a plank” to “soft as a cloud”)



# Ranked and measurement data

- Similarity: for every two ranks, the third between them has sense
- E.g., it is possible to imagine mattress with softness between 2 and 3
- However, ranks are not represent intervals correctly!
- Ranked data should be studied with non-parametric methods



# How to create ranked data

In R, ranked data is normally represented by the same numerical vector or *ordered factor*. Command `cut()` will break continuous data into ranks:

```
> height <- trees[,2]
> cut(height, 3, labels=c(1:3), ordered=T)
> cut(height, 3, ordered=T)
```



# Finishing...

## Save your commands!

`(savehistory(<today's date>.r)` or File -> Save as... on  
Mac)



# For Further Reading



A. Shipunov.

*Biometry* [Electronic resource].

2012—onwards.

Mode of access:

[http://ashipunov.info/shipunov/school/biol\\_240](http://ashipunov.info/shipunov/school/biol_240)



A. Shipunov, and others.

*Visual statistics. Use R!*

Ongoing translation from Russian.

