

Biometry. Lecture 19

Alexey Shipunov

Minot State University

April 18, 2016



- 1 2D statistics
 - Anatomy of relation: regression



```
> 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)



2D statistics

Anatomy of relation: regression



Blood data example

- 24 rows and 2 columns data for observations of ventricular velocity with different levels of blood glucose
- Data was taken from patients with diabetes type I.



Running the example and explaining results

```
> # install.packages("ISwR")
> library(ISwR)
> str(thuesen); head(thuesen)
> thuesen <- na.omit(thuesen)
> thuesen.lm <- lm(short.velocity ~ blood.glucose,
+ data=thuesen)
> thuesen.lm
> summary(thuesen.lm)
```



Scatterplot with regression line

```
> plot(short.velocity ~ blood.glucose, data=thuesen)  
> abline(thuesen.lm)
```



Visualizing residuals

```
> with(thuesen, segments(blood.glucose,  
+ fitted(thuesen.lm), blood.glucose, short.velocity))
```



Confidence intervals for regression

```
> source("http://ashipunov.info/r/asmisc.r")  
> plot(short.velocity ~ blood.glucose, data=thuesen)  
> abline(thuesen.lm)  
> Cladd(thuesen.lm, data=thuesen)
```



Regression diagnostics

```
> plot(thuesen.lm)
> # Compare with bad one:
> plot(lm(height ~ weight, data=women))
```



One more model

- We will try to understand the relation between gross state product (GDP) and rate of murders



One more model, analysis

```
> gdp2010 <- read.table("http://ashipunov.info/data/gdp2010.txt",
+ h=T, sep="\t")
> mg <- data.frame(murder=USArrests$Murder,
+ gdp=gdp2010$GDP)
> mg.lm <- lm(murder ~ gdp, data=mg)
> summary(mg.lm)
> plot(mg.lm) # NOT a plot of model: 4 diagnostic plots
> plot(murder ~ gdp, data=mg) # Plot of model
> abline(mg.lm)
> source("http://ashipunov.info/r/asmisc.r")
> Cladd(mg.lm, data=mg)
```



Finishing...

Save your commands!

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



Summary: most important commands

- `cor()`—calculates correlation coefficients
- `cor.test()`—run correlation tests
- `lm()`—estimate the linear regression



For Further Reading



A. Shipunov.

Biometry [Electronic resource].

2012—onwards.

Mode of access:

http://ashipunov.info/shipunov/school/biol_240



A. Shipunov, and many others.

Visual statistics. Use R!

2016—onwards.

Mode of access: http://ashipunov.info/shipunov/school/biol_240/en/visual_statistics.pdf

