Biometry. Lecture 23

Alexey Shipunov

Minot State University

May 5, 2014

Outline

Outline

Starting...

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

or

“Change dir”

in menu!

Questions and answers

Previous final question

Previous final question: the answer

What are the null and alternative hypotheses for ANOVA?

Previous final question: the answer

What are the null and alternative hypotheses for ANOVA?

- Null: no one is different from all others
- Alternative: at least one is different from all others

Once more about logistic regression

```
> p <-
+ read.table("http://ashipunov.info/data/plantago_samples.txt",
+ sep="\t", h=T)
> ...
```

Is the success of PCR related with age of sample?

Once more about ANOVA

```
> p <- read.table("http://ashipunov.info/data/molds.txt",
+ sep="\t", h=T)
> ...
```

Does the diameter of colony differ in different mold species?

Multivariate statistics, or Data Mining

Generic methods

Multivariate statistics, or Data Mining

- Multivariate statistics is mostly the **analysis of structure**
- Majority of multivariate methods are just visualization via reduction of dimensions (projection)
- Inferential methods also exist

Generic methods

- These methods simply take into account more than two variables
- Conditional (trellis) plots and 3D cubes and surfaces are most common representatives of this group in R

Trellis plots

From now on, we will frequently use embedded "`iris`" data from Fisher.

```
> coplot(Sepal.Length ~ Petal.Length | Species, data=iris)
> library(lattice)
> xyplot(Sepal.Length ~ Petal.Length + Petal.Width
+ | Species, data=iris)
```

Matrix graph

```
> pairs(iris[1:4], pch=21, bg=(1:3)[iris$Species])
```

Pictographs

```
> stars(mtcars[1:9,1:7])
# install.packages("TeachingDemos") if you do not have it
> library(TeachingDemos)
> faces(mtcars[1:9,1:7])
```

`mtcars` is a data about different design and performance of 32 cars (1973–74 models) like: (1) mpg, (2) Number of cylinders, (3) Displacement, (4) Gross horsepower, (5) Rear axle ratio, (6) Weight (lb/1000) (7) qsec 1/4 mile time.

Finishing...

```
> savehistory("20140505.r")
```

Final question (2 points)

Final question (2 points)

Why do biologists need multivariate analysis?

Summary

- `coplot()`—make a conditional plot

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. [Under translation from Russian.]