Introduction to UsingR

Wim Krijnen
Lector Analyse Technieken voor Praktijkonderzoek
Lectoraat Healthy Ageing, Allied Health Care and Nursing
Hanze University of Applied Sciences

July 1, 2015

Outline

Purpose of this course

- Open for anybody against minimal costs
- Give brief introduction in using R
- Getting you started and self sufficient
- Providing help, examples, tutorials, (free) literature
- Solving some of your problems
- Continuation on demand, once in 3 months
- Starting a community: share ideas and helping each other

Outline

How to do this course

- Bring in your own work and start from there
- Study and learn from examples
- Ability to write correct scripts determines usefulness
- Come up with ideas, problems, mistakes, and frustrations!
- Be active and explore!
- Step in the world of statistical programming
- Enjoy the power of programming

Outline

Situations for Using R

- Repeated similar problems
- Programming of visualizations: publication ready plots
- Handle large data sets
- Desire flexibility in statistical programming
- Use of modern techniques (bootstrap, robust, Baysian)

Outline

Reasons for Using R

- Widely used in statistics and applied sciences
- Reliable free open source
- Versatile: SPSS, Matlab, MySQL, Perl, JAVA, C++, Fortran
- Extensive help
- Numerous libraries with modern methods
- High level language with many built-in-functions

Disadvantages:

- Steep learning curve (use it regularly)
- Command line, some GUI
- Not fastest (C++)

Outline

Some advise on starting

- Analyze needs and whether it fits to R
- R for Beginners
- Simple R
- IcebreakR
- Work on your own problems
-
-

Outline

Install R

URL: `http://cran.r-project.org`
choose operating system: Windows, Linux, Mac
choose `base`
html help

installing libraries

Install a library `chooseCRANmirror()`

```
install.packages(c("TeachingDemos"),repo=
"http://cran.r-project.org",dep=TRUE)
library(TeachingDemos)
plot(dice(12,1))
```

Install a bundle:

```
install.packages("ctv")
library("ctv")
install.views("Robust")
install.views("Psychometrics")
install.views("Econometrics")
install.views("SocialSciences")
```

Outline

Well Written Manuals

- An Introduction to R
- The R Language Definition
- R Installation and Administration
- R Data Import/Export

Searching and Tutorials

From `http://cran.r-project.org`

- `Search`, `Task Views`, `Manuals`, `FAQs`, `The R Journal`, `Wiki`
- → `Contributed`: "R for Beginners", "The R Guide", "Practical Regression and Anova using R", "IcebreakeR", "Notes on the use of R for psychology experiments and questionnaires", "An Introduction to S and The Hmisc and Design Libraries"
- → `Contributed` "R reference card",

Useful tutorial for beginners:
`http://cran.r-project.org/doc/contrib/Verzani-SimpleR.pdf`
`http://www.r-project.org` → choose `Books`

Help on examples

from commandline

- `help.start()`
- `library()`, `ls(package:stats)`, `library(help="stats")`
- `help(t.test)`, `?sum`, `??solve`, `apropos("if")`
- `methods(plot)` plotting functions
- `example(boxplot)` examples
- `demo()` demonstrations of code
- `mean.default` study code of function

Help on Programming

- help(Control) : “for” and “ while” loops
- help(Syntax) : syntax of operators
- help(Logic) : logical operators AND, OR, negation
- help(Arith) : on arithmetic, relational, logical operators, mathematical functions
- help(Special) : gamma function

Outline

Reading and writing

Reading		Writing	
function	library	function	library
scan	base	write.table	base
read.table	utils	write.table	base
read.csv	utils	write.csv	utils
source	base	save	base
read.spss	foreign		
		x11	grDevices
		postscript	grDevices
		xtable	xtable (LaTeXusers)

Outline

Important Functions

scalar, vector, matrix, list, function, environment;
Everything is an object, belonging to a class!
Once defined extract information from it, use functions on it.
`q(),history`; quit, previous commands
`rm(),rm(list=ls())`; remove objects
`ls(),objects()`; listing of objects
`class(x),str(x)`; class or structure of object x
`getwd,setwd,dir`; get set working directory, interaction with OS
`numeric, character, matrix,data.frame,list`; construct object
`factor,gl`; construct factor
`summary,residuals,coef`; generic functions for lm, glm etc.
`mean,median,sd,IQR,quantile`; descriptive statistics
`plot,matplot`; plotting
`rownames,colnames,rowcolnames`
`function,apply,lapply`; apply function on row/columns of matrix
`grep,regexpr`; regular expressions

Outline

Some Plots of data

Graphical representation	R function
Box-and-Wiskers	`boxplot`
Histogram	`hist`
Pie	`pie`
Bar	`barplot`
Density	`plot(density())`
Dot Plot	`stripchart`

Outline

Some Statistical Tests

Test	H_0	function
t-test	$\mu 1 = \mu$, $\mu_1 = \mu_2$	`t.test`
Wilcoxon; Mann-Whitney	$F = G$	`wilcox.test`
ANOVA	$\mu_1 = \mu_2 = \mu_3$	`anova`
Kruskal-Wallis	$F_1 = F_2 = F_3$	`kruskal.test`
association	$\tau = \tau_0$	`cor.test`
probability of success	$p = p_0$	`binom.test`
normality	X norm. distrib.	`shapiro.test`
Kolmogorov-Smirnov	$F = G$	`ks.test`
F-test	$\sigma_1 = \sigma_2$	`var.test`
Fisher's exact test	independence	`fisher.test`

For loop and apply

```
apply(matrix,margin,fun)
col.means <- apply(x,2,mean)
col.means <- numeric
for (j in 1:ncol(x))
col.means[j] <- mean(x[,j])
```

Example: Daily energy intake

Daily energy intake (Altman, 1991, p.183) of group of woman; recommended intake 7725 kJ

$H_0 : \mu = 7725$kJ, $H_0 : \mu \neq 7725$kJ

```
> x <- c(5260,5470,5640,6180,6390,6515,6805,7515,7515,
> t.test(x, mu=7725)

        One Sample t-test

data:  x
t = -2.8208, df = 10, p-value = 0.01814
alternative hypothesis: true mean is not equal to 7725
95 percent confidence interval:
 5986.348 7520.925
sample estimates:
mean of x
 6753.636
```

Conclusion: H_0 not rejected