




Bayesian Statistics and Hierarchical Bayesian Modeling for Psychological Science

Lecture 01

Lei Zhang

Social, Cognitive and Affective Neuroscience Unit (SCAN-Unit)
Department of Basic Psychological Research and Research Methods

https://github.com/lei-zhang/BayesCog_Wien

lei.zhang@univie.ac.at
lei-zhang.net
 @lei_zhang_lz

Goal of this course

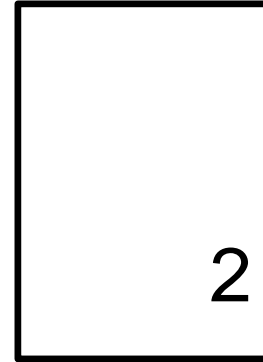
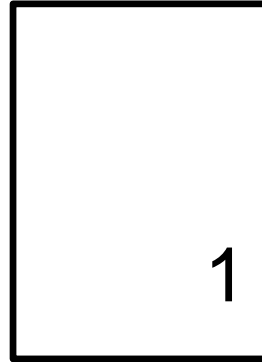
- Practical R programming
- Practical model-building in Stan, model diagnostics
- (Enough) theory to ground you
- Be comfortable to use R/Stan for your own work



Schedule of Lectures

2w	[N	Tuesday 05.03.	L01	Overview; Introduction to R	Announce programming project
			Tuesday 19.03.	L02	Probability; Bayes' Theorem	
			Tuesday 26.03.	L03	Binomial model; MCMC & Stan	
			Tuesday 02.04.	L04	Simple linear model	
3w	[Tuesday 09.04.	L05	Cognitive Modeling; RL model	Review of a paper: #1
			Tuesday 30.04.	L06	More on RL model	
			Tuesday 07.05.	L07	Hierarchical modeling	
			Tuesday 14.05.	L08	More on hierarchical modeling	
			Tuesday 21.05.	L09	Optimizing Stan codes	
			Tuesday 28.05.	L10	PRL task & model comparison	
2w	[Tuesday 04.06.	L11	Introduction to model-based fMRI	Review of a paper: #2
			Tuesday 18.06.	L12	Stan style tip & debugging	
			Tuesday 25.06.	L13	In-class quiz; HPC demo	Quiz

Review of a paper?



After L05

students 1:10

students 11:20

After L11

students 11:20

students 1:10

How to review a paper?

- Suppose you are invited by a journal editor to review a paper
- Of course, you have to read it 😊, carefully and critically
- Then write a review report to the editor
 - (1) Make a summary. What is this paper about? What was done? What was the conclusion?
 - (2) List your concerns. Is the design appropriate? Are the analyses sound? Do their data support the conclusion? What can be done better?
- For this course:
 - up to 3 pages (11 pt, 1.5 space)
 - be independent: okay to discuss HOW to review, but do NOT discuss WHAT to review

Programming project

- will be announced after L02
- can be submitted at any time before end of semester (30.06)
- use R and RStan
- will be a real-world cognitive modeling problem
- hand in the *.R and *.stan files in a ZIP file
- name as: lastname_matriculatenummer_200075.ZIP
- no need to write a report

Quiz?

- In-class quiz on L13
 - 10 multiple choices
 - 30min
 - Feel free to use R
 - But NOT Google!
-
- We will discuss solutions afterwards



KEEP
CALM
ITS
QUIZ

Gradings

- Regular participation (20%)
 - Review of paper#1, 10 (25%), due on 30.04.2019*
 - Review of paper#2, 10 (25%), due on 25.06.2019*
 - Programming work, 10 (20%), announce after L02, due on 30.06.2019*
 - Quiz 10 (10%), in-class on 25.06.2019
-
- Grades: >90% 1, >80% 2, >70% 3, >60% 4, <60% 5
 - At least 60% to obtain 4 ECTS

*0.5 point deduction per day for overdue submission

Overview

What is your experience with...

- Statistics?
- R? (and / or Matlab?)
- Cognitive Modeling?

You would like to...

- gain knowledge of Bayesian stats?
- be able to read “computational modeling” section in papers?
- write your own model?

Survey results

sent to 27 people (registered + waiting list)

received 24

89% receive rate, many thanks!

spontaneous feedback are still more than welcome!

Overview

This course is NOT about...

- ... Bayes in the brain (e.g. predictive coding)
- ... Bayesian statistics to supersede classic statistics



180160 SE New Trends in Cognitive Science - The predictive coding approach to mind/cognition

However, Bayesian statistics offer great tools to analyze **cognitive processes**!

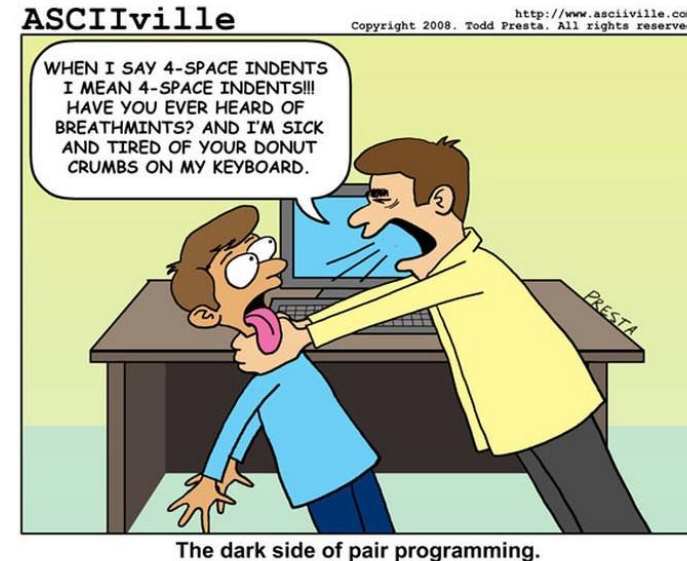
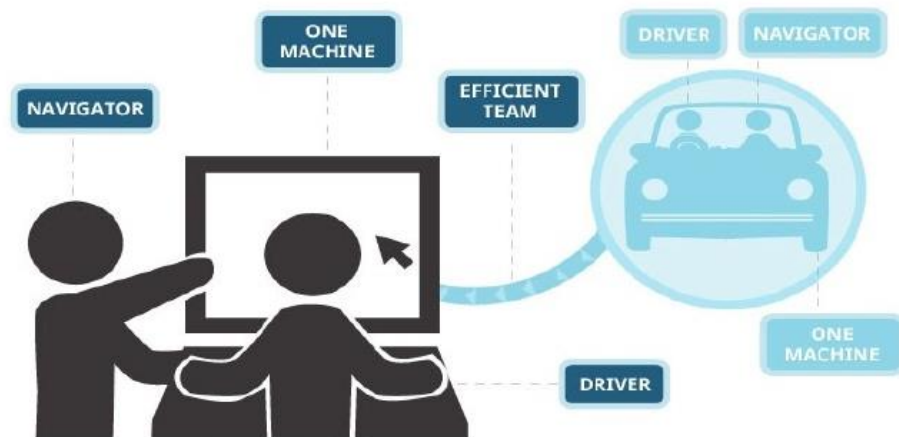
- Construct cognitive models
- Estimate posterior distributions of parameters
- Compare models: which is the best one, given the data
- Perform model-based analysis, e.g. model-based fMRI/EEG/eye-movement

A photograph of a chalkboard with the Bayesian formula $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$ written in blue chalk. The formula is written in a slightly messy, handwritten style. The background is dark, and the chalk is bright blue.

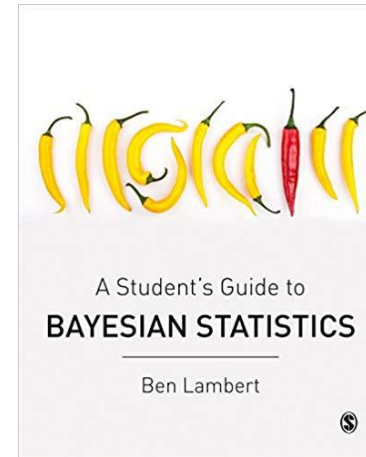
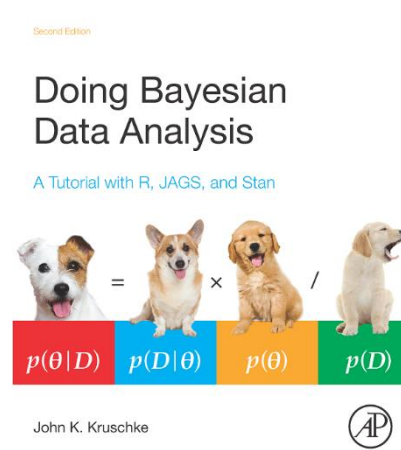
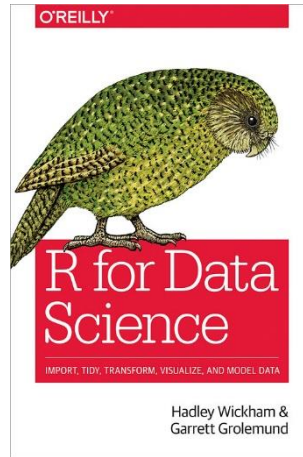
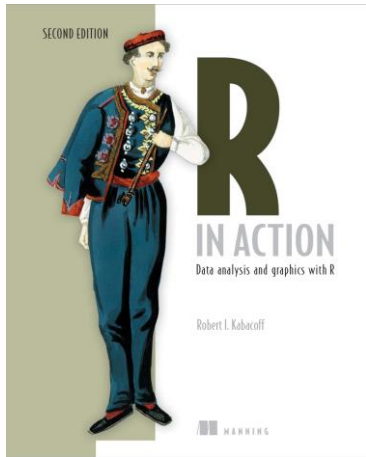
How to Get the Most out of the course

- Lecture structure: 60min theory + demo, 20-30min exercise + discussion
- Work in pairs: Talk to each other & help each other
- Ask questions
- Try the exercises

PAIR PROGRAMMING



Resources

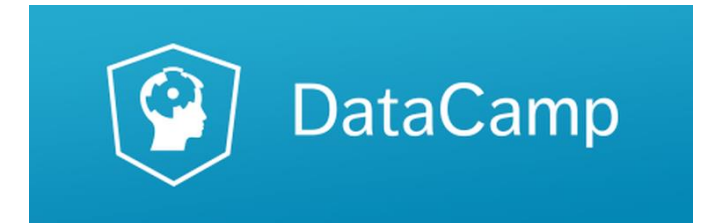


Statistical Thinking for the 21st Century

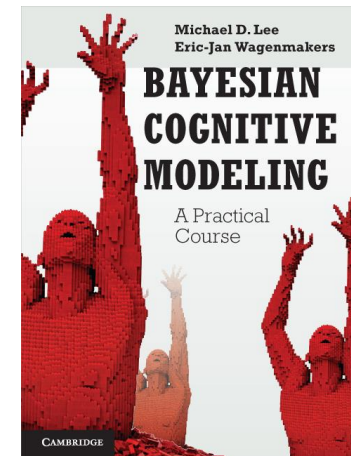
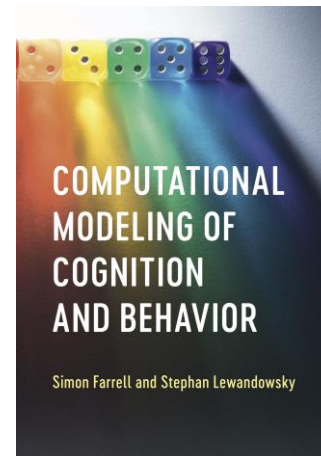
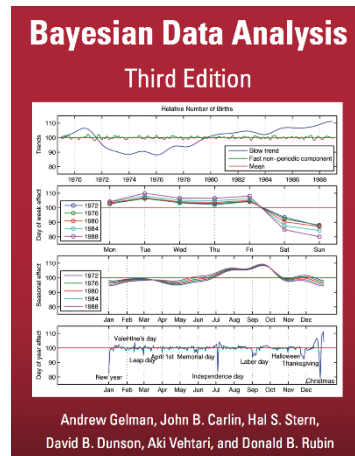
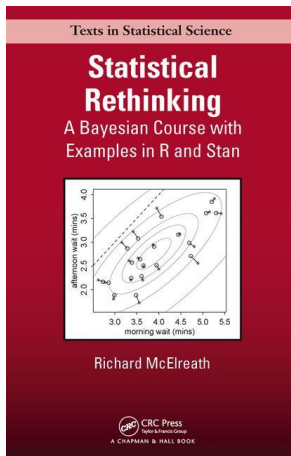
Copyright 2018 Russell A. Poldrack

Draft: 2018-11-22

<http://thinkstats.org/>



<https://www.datacamp.com/>



<https://jasp-stats.org/>

BASICS OF R PROGRAMMING



R Basics

cognitive model

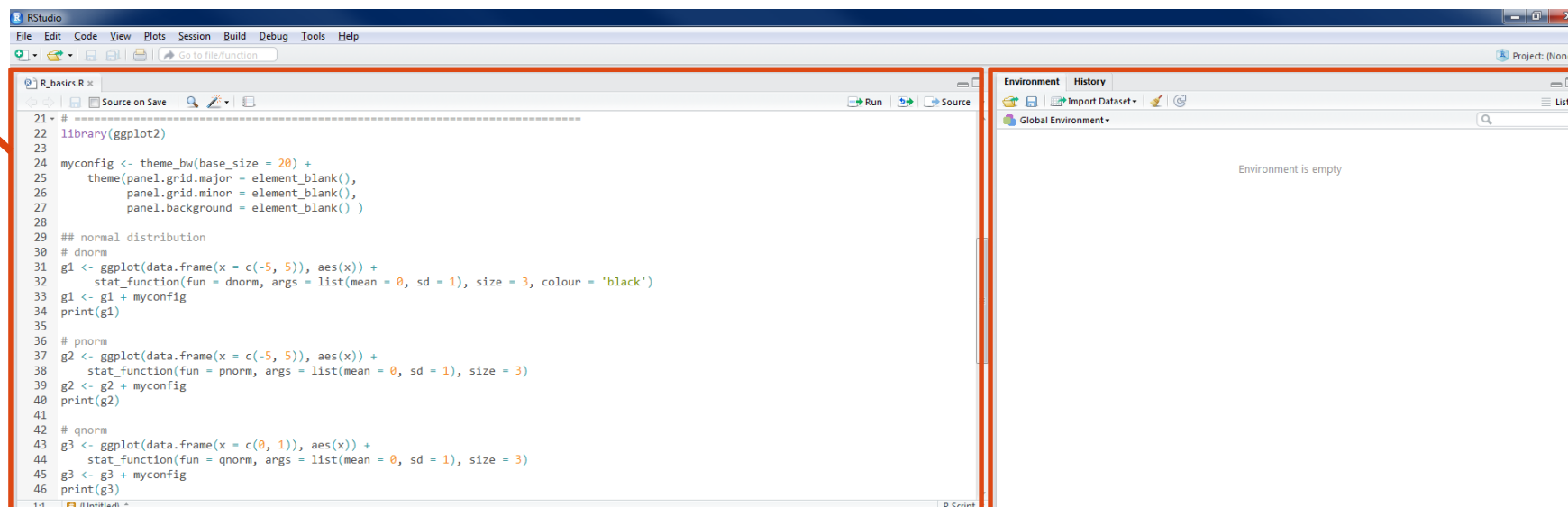
statistics

computing

- R
 - a programming language for statistical computing
 - R has its own user interface
 - freely available on Windows, Mac, and Linux
- R Studio
 - integrated development environment (IDE) for R
 - a more sophisticated R-friendly editor, with helpful syntax highlight



script editor

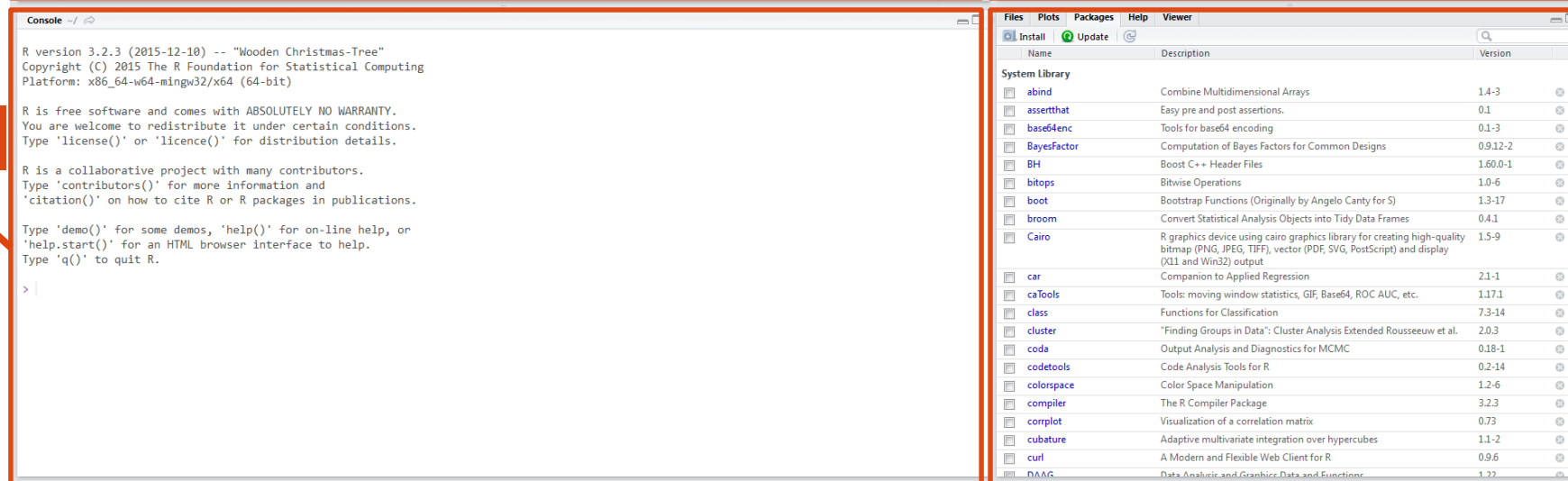


The image shows the RStudio interface. The top-left pane is the script editor, displaying R code for creating a ggplot2 theme and three plots (g1, g2, g3). The top-right pane is the Environment/History pane, which is currently empty, showing 'Global Environment' and 'Environment is empty'.

```
21 #  
22 library(ggplot2)  
23  
24 myconfig <- theme_bw(base_size = 20) +  
25   theme(panel.grid.major = element_blank(),  
26         panel.grid.minor = element_blank(),  
27         panel.background = element_blank())  
28  
29 ## normal distribution  
30 # dnorm  
31 g1 <- ggplot(data.frame(x = c(-5, 5)), aes(x)) +  
32   stat_function(fun = dnorm, args = list(mean = 0, sd = 1), size = 3, colour = 'black')  
33 g1 <- g1 + myconfig  
34 print(g1)  
35  
36 # pnorm  
37 g2 <- ggplot(data.frame(x = c(-5, 5)), aes(x)) +  
38   stat_function(fun = pnorm, args = list(mean = 0, sd = 1), size = 3)  
39 g2 <- g2 + myconfig  
40 print(g2)  
41  
42 # qnorm  
43 g3 <- ggplot(data.frame(x = c(0, 1)), aes(x)) +  
44   stat_function(fun = qnorm, args = list(mean = 0, sd = 1), size = 3)  
45 g3 <- g3 + myconfig  
46 print(g3)
```

environment/
command history

console



The image shows the RStudio interface. The bottom-left pane is the console, displaying the R version (3.2.3), copyright information, and instructions for using R. The bottom-right pane is the Packages pane, showing a list of installed and available packages with their descriptions and versions.

R version 3.2.3 (2015-12-10) -- "Wooden Christmas-Tree"
Copyright (C) 2015 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |

Name	Description	Version
System Library		
abind	Combine Multidimensional Arrays	1.4-3
assertthat	Easy pre and post assertions.	0.1
base64enc	Tools for base64 encoding	0.1-3
BayesFactor	Computation of Bayes Factors for Common Designs	0.9.12-2
BH	Boost C++ Header Files	1.60.0-1
bitops	Bitwise Operations	1.0-6
boot	Bootstrap Functions (Originally by Angelo Canty for S)	1.3-17
broom	Convert Statistical Analysis Objects into Tidy Data Frames	0.4.1
Cairo	R graphics device using cairo graphics library for creating high-quality bitmap (PNG, JPEG, TIFF), vector (PDF, SVG, PostScript) and display (X11 and Win32) output	1.5-9
car	Companion to Applied Regression	2.1-1
caTools	Tools: moving window statistics, GIF, Base64, ROC AUC, etc.	1.17.1
class	Functions for Classification	7.3-14
cluster	"Finding Groups in Data": Cluster Analysis Extended Rousseeuw et al.	2.0.3
coda	Output Analysis and Diagnostics for MCMC	0.18-1
codetools	Code Analysis Tools for R	0.2-14
colorspace	Color Space Manipulation	1.2-6
compiler	The R Compiler Package	3.2.3
corplot	Visualization of a correlation matrix	0.73
cubature	Adaptive multivariate integration over hypercubes	1.1-2
curl	A Modern and Flexible Web Client for R	0.9.6
DAAG	Data Analysis and Graphics: Data and Functions	1.22

file/pkg/img/
etc.

Know your R

```
>R.version
```

```
platform      _  
arch           x86_64-w64-mingw32  
os            mingw32  
system        x86_64, mingw32  
status  
major          3  
minor          5.1  
year          2018  
month         07  
day           02  
svn rev       74947  
language       R  
version.string R version 3.5.1 (2018-07-02)  
nickname      Feather Spray
```

R Console as a Calculator

cognitive model

statistics

computing

Addition and Subtraction

```
> 3+2  
[1] 5
```

```
> 3-2  
[1] 1
```

Multiplication and Division

```
> 3*2  
[1] 6
```

```
> 3/2  
[1] 1.5
```

Exponents in R

```
> 3^2  
[1] 9
```

```
> 2^3  
[1] 8
```

Constants in R

```
> pi  
[1] 3.141593
```

```
> exp(1)    base of the natural logarithm  
[1] 2.718282
```

Special values

Infinite Values

```
> Inf  
[1] Inf
```

```
> 1+Inf  
[1] Inf
```

Machine Epsilon

```
> .Machine$double.eps  
[1] 2.220446e-16
```

```
> 0>.Machine$double.eps  
[1] FALSE
```

Empty Values

```
> NULL  
NULL
```

```
> 1+NULL  
numeric(0)
```

Missing Values

```
> NA  
[1] NA
```

```
> 1+NA  
[1] NA
```

Storing and manipulating variables

Define objects `x` and `y` with values of 3 and 2, respectively:

```
> x=3
```

```
> y=2
```

Some calculations with the defined objects `x` and `y`:

```
> x+y
```

```
[1] 5
```

```
> x*y
```

```
[1] 6
```

Warning: R is case sensitive, so `x` and `X` are not the same object.

Basic R functions

cognitive model

statistics

computing

Combine

```
> c(1, 3, -2)
[1] 1 3 -2
```

```
> c("a", "a", "b", "b", "a")
[1] "a" "a" "b" "b" "a"
```

Sum and Mean

```
> sum(c(1, 3, -2))
[1] 2
```

```
> mean(c(1, 3, -2))
[1] 0.6666667
```

Variance and Std. Dev.

```
> var(c(1, 3, -2))
[1] 6.333333
```

```
> sd(c(1, 3, -2))
[1] 2.516611
```

Minimum and Maximum

```
> min(c(1, 3, -2))
[1] -2
```

```
> max(c(1, 3, -2))
[1] 3
```

Basic R functions (cont.)

Define objects `x` and `y`:

```
> x=c(1,3,4,6,8)
```

```
> y=c(2,3,5,7,9)
```

Calculate the correlation:

```
> cor(x,y)
```

```
[1] 0.988765
```

Calculate the covariance:

```
> cov(x,y)
```

```
[1] 7.65
```

Combine as columns

```
> cbind(x,y)
```

	x	y
[1,]	1	2
[2,]	3	3
[3,]	4	5
[4,]	6	7
[5,]	8	9

Combine as rows

```
> rbind(x,y)
```

	[, 1]	[, 2]	[, 3]	[, 4]	[, 5]
x	1	3	4	6	8
y	2	3	5	7	9

Basic Commands

```
getwd()
setwd('E:/teaching/BayesCog_Wien/')
dir() # folders/files in the wd
ls()  # anything in the environment/workspace
print('Hello World!')
cat('Hello', 'World!')
paste0('C:/', 'Group1')
help(func)
? func # and Google!
a <- 5
a = 5
head(d) # first 6 entries
tail(d) # last 6 entries
save(varname, file = "pathname/varname.RData")
load("pathname/varname.RData")
rm(list = ls())
q()
```

RStudio - Shortcuts

cognitive model

statistics

computing

Ctrl + L: clean console

Ctrl + Shift + N: create a new script

↑: command history

Ctrl(hold) + ↑: command history with certain starts

Ctrl + Enter: execute selected codes (in a script)

Editor (WIN general) - Shortcuts

cognitive model

statistics

computing

Ctrl + home/Pos: go to the very top of a script

Ctrl + end/Ende: go to the very end of a script

Shift(hold) + ↑/↓: select line(s)

Ctrl(hold) + ←/→: select word(s)

Data Classes

cognitive model

statistics

computing

numeric: 1.1 2.0

integer: 1 2 3

character / string: "hello world!"

logical: TRUE FALSE

factors: "male" / "female"

(complex: 1+2i)

Data Types

cognitive model

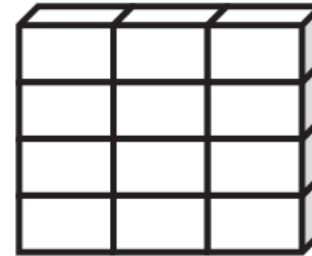
statistics

computing

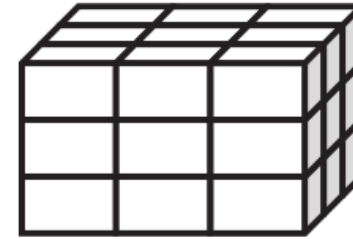
(a) Vector



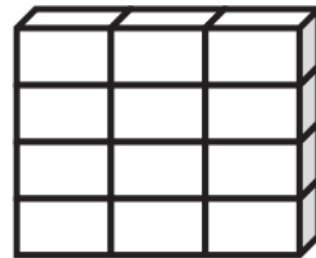
(b) Matrix



(c) Array

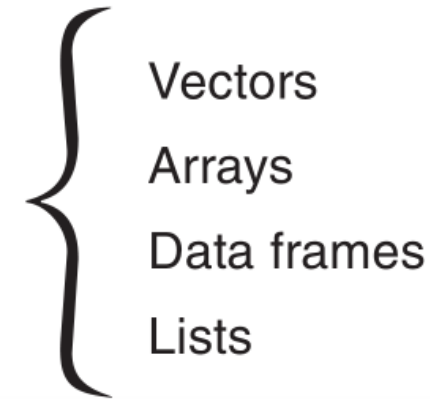


(d) Data frame



Columns can be different modes

(e) List



Exercise I

cognitive model

statistics

computing

```
.../01.R_basics/_scripts/R_basics.R
```

up to “Control Flow”

TASK: practise basic R commands and data type

TIP: `class()`, `str()`

Side note: folder structure

