

Exercises for the course
“An introduction to R”
 Sheet 01

Exercise 1: If R is not yet installed, then install it now, see the lecture notes. Use R to calculate

$$3^7, \quad \binom{22}{17}, \quad 8!, \quad \sqrt{\pi}.$$

Exercise 2: *Vectors are used for everything in R. So handling vectors is the first thing to learn.*

- Define the variable **v1** as the vector

$$(3, 7, -4, 0)$$

View the vector by entering **v1** on the R command line.

- Define the variable **v2** as the vector

$$(1, 2, 3, \dots, 48, 49, 50)$$

- Define the variable **v3** as the vector

$$(3, 7, -4, 0, 1, 2, 3, \dots, 48, 49, 50)$$

You may use **v1** and **v2** for this.

- Define the variable **v4** as the vector

$$(0.0, 0.1, 0.2, 0.3, \dots, 1.8, 1.9, 2.0)$$

- Define the variable **treatment** as a vector of length 100 with elements

$$("yes", "control", "yes", "control", \dots, "yes", "control")$$

- Define the variable **smoker** as a vector of length 99 with elements

$$("yes", "no", "no", "yes", "no", "no", \dots, "yes", "no", "no")$$

- Sum over all elements of **v1**. Sum over all elements of **v2**.

- What is the product of all elements of the vector

$$(10, 11, 12, 13, \dots, 19, 20)?$$

Exercise 3: We assume the vectors `smoker` and `treatment` of Exercise 2 to be known. If you didn't do Exercise 2, then write down the respective R commands which answer the following questions. Define the vector

```
lifespan <- abs( round( 100*sin(1:99) ) )
```

- Suppose we have 99 individuals of which we have measured some quantity and stored in the vector `lifespan`. You may think of `lifespan` as the life span of the individuals but keep in mind that the data is self-generated and not meaningful. The vector `smoker` tells us which individual is smoker and which is non-smoker. Now we wish to study the measured quantity of all smokers. Define a new vector `x` which consists of all elements of `lifespan` at whose index in `smoker` is the element "yes". What is the maximum of `lifespan` over all smokers?
- Define a new vector `y` of all even elements in `lifespan` which are greater than or equal to 16. What is the minimum of all these elements?
- Half of the individuals got a certain treatment. Produce a new vector consisting of the lifespans of all individuals which are smokers and got the treatment.
- Produce a new vector of the lifespans of individuals which are non-smokers or got the treatment ('or' is not exclusive).

Exercise 4: *It produces faster code and needs less typing to use `sum()` and `prod()` instead of using loops.*

Using the commands `sum()` and `prod()`, calculate

$$\sum_{i=30}^{200} i, \quad \sum_{i=1}^{100} \frac{1}{i}, \quad \sum_{i=0}^{100} i * e^{-i}, \quad \prod_{i=1}^{100} (2 \cdot i^2 - i)$$

Exercise 5: *The most important R command is `help()`. It is good to get used to it as soon as possible.*

The commands `signif()` and `expm1()` have not been discussed in the course. Use `help()` (or `?`) to learn how to use them.

- Produce a vector which contains each element of the vector $(1 : 100)^8$ rounded to 3 significant digits. Recall that e.g. 42598 rounded to 3 significant digits is $4.26 * 10^4$.
- For each element x_i of the vector $(10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, \dots, 10^{-17}, 10^{-18})$ calculate

$$e^{x_i} - 1$$

first by using the R command `exp()` and then by using the R command `expm1()`. Which result do you trust more?

Exercise 6: *This exercise is supposed to give you an impression of R*

Download the files 'R.intro.R', 'FinchesSulloway.txt', 'swarth.dat' from the course web page. Open a graphical user interface for R (e.g. RGUI.exe for Windows, R.app for Mac, Linux: start R, type `install.packages("Rcmdr",dependencies=TRUE)`, then type `library(Rcmdr)`, this opens a graphical user interface). Then click on 'open script' in the menu and open 'R.intro.R'. Now execute the lines of that script line by line and see what happens. Go through that script until 'Abschnitt 14'. The explanations are in german but you can still watch what happens.