

## Exam for Bachelor / Master-Module *Computational Statistics with R*

00.00.0000

0:00 p.m. - 0:00 p.m.

|                     |                                 |
|---------------------|---------------------------------|
| Course:             | Computational Statistics with R |
| Contact Hours:      | 3 SWS                           |
| Duration:           | 45 minutes                      |
| Examiner:           | Univ.-Prof. Dr. Ulrich Küsters  |
| Allowed Appliances: | Calculator                      |

### Test Exam

Before you start, check if your exam is complete. This exam consists of 3 exercises and 6 pages. **All exercises** have to be solved. You may answer in **German or Englisch**, but you have to decide for either one of them. If your answers are not consistent in either German or Englisch, the language of your first answer will determine your chosen language.

In case one (sub-) task cannot be answered due to incomplete or inconsistent specifications, give a **short but precise** explanation why the (sub-) task is not solvable. You will then get full marks for the concerning (sub-) task.

If you are allowed to use a formulary, this means exclusively the formulary of the Chair of Statistics without additional highlighting or additional marks. Own writing paper is not allowed.

Please round results to **four** decimal places after the comma. Only for exercises with a clear and reproducible approach, points will be awarded. Write down your solutions directly under the questions. If you run out of space, please use the back of the previous page. Additionally, there are 1 auxiliary pages at the end of the exam for your answers.

Good luck!

Chair of Statistics

|                 |                     |
|-----------------|---------------------|
| wws-Identifier: | Immatriculation No: |
|-----------------|---------------------|

# Evaluation Sheet

| EXCERCISE | MAXIMUM<br>POINTS | OBTAINED<br>POINTS |
|-----------|-------------------|--------------------|
| 1         | 20                |                    |
| 2         | 20                |                    |
| 3         | 10                |                    |
| SUMME     | 50                |                    |

Please hand this excercise sheet back in  
with your exam solutions

## 1 Exercise (20 Points)

Explain the following concepts, terms or functions, including their benefits and/or drawbacks:

- a) Coercion
- b) Integrated Development Environment

Compare and briefly discuss the following features, functions or reserved constants:

- c) `for` and `while`
- d) `NA`, `NULL` and `NaN`

## 2 Exercise (20 Points)

The following code was found in the file `project.final.final2.R` as part of an old undocumented project. Therefore its purpose is unknown. Being a young and ambitious R developer you were asked to answer the following questions.

```
R> rows <- 5
R> cols <- 5
R>
R> mat <- matrix(FALSE, nrow = 5, ncol = 5)
R>
R> for (idx in 1:rows){
R>   for (jdx in 1:cols){
R>     if ( idx < jdx ){
R>       mat[idx, jdx] <- 1
R>     } else if ( idx = jdx ){
R>       mat[idx, jdx] <- 1
R>     }
R>   }
R> }
R>
R>
R> vec <- apply(mat, 1, FUN=sum, na.rm = T)
R>
R> vec
```

- a) While executing the code above the following message appears. Explain why the message is returned and correct the error in the provided source code.

```
Error: unexpected '=' in:
"      mat[idx, jdx] <- 1
} else if ( idx ="
```

- b) Make yourself familiar with the provided code and briefly describe in a generic way what the code does (line-by-line, at least one full iteration of the outer loop). Ensure that your description is short but precise and allows the reader to gain a deep understanding of the data manipulation process.
- c) Write down the content of the variable `vec` in the last line **after** fixing the error and executing the code above.

### 3 Exercise (10 Points)

The following excerpt gives an overview of an object containing information about a typical American family. Please provide answers to the following questions:

```
R> family.data
```

|   | ID | Name  | Age | Sex | Height | Weight |
|---|----|-------|-----|-----|--------|--------|
| 1 | 1  | Homer | 38  | m   | 182    | 108    |
| 2 | 2  | Marge | 34  | f   | 223    | 58     |
| 3 | 3  | Bart  | 10  | m   | 122    | 35     |
| 4 | 4  | Lisa  | 8   | f   | 120    | 33     |

```
R> str(family.data)
```

```
'data.frame':  4 obs. of  6 variables:
 $ ID      : int  1 2 3 4
 $ Name    : Factor w/ 4 levels "Bart","Homer",...: 2 4 1 3
 $ Age     : num  38 34 10 8
 $ Sex     : Factor w/ 2 levels "f","m": 2 1 2 1
 $ Height: num  182 223 122 120
 $ Weight: num  108 58 35 33
```

- What are the reasons that this object was created as a `data.frame`?
- How can R be used to return the numbers of rows and columns? Please make a suggestion for a function and write down the returned values.
- How can R be used to calculate the arithmetic mean of the variables `Height`, `Weight` and `Age`? Please comment on your provided solution in terms of efficiency.

**Auxiliary page for your answers**  
**Please mark clearly to which exercise your answer belongs.**