Statistical Computing Master Data Science



Winter Semester 2017/18

Prof. Tim Downie

R Test Example version December 2017

(Time allowed: 60 minutes)

Name, Forename:	MatrikeInr.:
	ven boxes. If your answer is too long to fit in the box, please separate A4 paper, indicating clearly the question number

Preliminaries:

You need to download a file which contains the following R-objects. Reciepts, datf and datf2

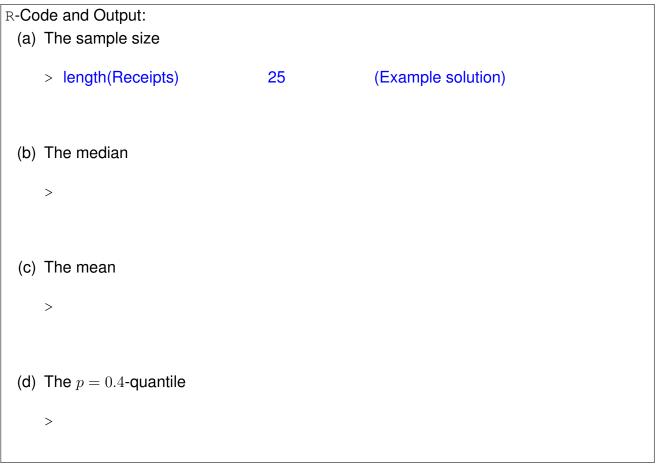
- Download the file noting which directory the file is stored in.
- Start RStudio.
- If required set your working directory.
- load the Data using
 - > load("TestOneExample.Rda")
- As a check type
 - > loaded()

You should get the text "The data for the test have loaded properly."

Exercise 1 (5 points)

The R object Receipts is a sample of till receipt totals (in Euros) from a café-Bar.

Enter the R Commands and the results to obtain the relevant statistics from these data.



- (e) Give the command to obtain a histogram of the receipts. To get maximum points your histogram should include the following properties:
 - The interval breaks should be at €0, 5, 10, ..., 50 located.
 - The histogram columns should be coloured grey
 - A sensible title and axis labels.

R-Co	ode:
>	
	se 2 (5 points)
	ve downloaded and installed two data frames called datf and datf2.
1	de and Output)
(a) (One command to give the number of rows and variables in datf
	>
(b) A	A command which outputs the variable names in datf
()	
(c) A	A frequency table for the variable in datf called ddd
	>
(d) A	A box-plot for aaa with a box for each level in ddd
(a) /	Took plot for alaa with a box for each level in add
	>
(e) -	The mean value of aaa for each level of ddd
	>
1	

	(D_(Coc	حا	an	Ы	\bigcirc	ıtn	ı ı + \
١	(K-1	-00	ıe	an	ıu	Οl	มเม	uı)

(f) A command which appends the columns of datf2 to the dataframe datf. The output is not required.

>

(g) What property of datf and datf2 is required for the command in part (f)

Exercise 3 (4 points)

(R-Code)

(a) Enter the commands and write the output:

```
> set.seed(100)
```

> rnorm(1)

(b) Generate a sample of 30 random numbers with a Normal distribution with mean 10 and standard deviation 5. Assign the output to an object called randsamp.

>

(c) Perform a one sample t-test testing whether the population mean is equal to 10.

>

(d) From the output of this command give the sample mean and the p-value.

>

Finishing off (1pt)

- Add a comment to the top of your script file with your name and matriculation number.
- Save your script file.
- Go to the section *Test 1 script files* in Moodle and upload your file. If there are any problems with this then please ask the lecturer.