Practical Exam

Sonja Hartnack, Terence Odoch & Muriel Buri

October 2018

Exercise 1

(a) Open a new R-Script and read in the `lepto_exam.csv` data file. This study data present a serology survey of leptospira sero-prevalence in rural and urban areas. Within the data set you find the following variables:

- `gender`: factor variabe with two levels "female" and "male"
- `age`: integer variable (years)
- `exposure`: factor variabe with two levels "urban" and "rural"
- `antibodies`: factor variabe with two levels "absent" and "presence"
- `num.rats`: integer variable (amount of rats seen in the last 3 months)

(b) Save the `lepto_exam` data frame as `lepto` data frame.

```
lepto <- data.frame(lepto_exam)
head(lepto)
str(lepto)
```

(c) Have a look at the `str(...)`, `summary(...)`, `head(...)`.

(d) How many observations are in the `lepto` data frame?

(e) Plot a boxplot of the `age` for each female and male (`gender`).

(f) Make a 2-by-2 table for `exposure` and `antibodies`.

(g) Evaluate with the help of a Chi-square test (`chisq.test(...)`) if the two categorical variables `exposure` and `antibodies` are related in any way. What's your conclusion?