

Exercises

Descriptive analysis and basic statistics in biomedical studies using R and Markdown

Juan R Gonzalez

juanr.gonzalez@isglobal.org

IACS - Instituto Aragonés de Ciencias de la Salud

Zaragoza, February 26th

Contents

1	R basics	1
2	Descriptive analyses	1

1 R basics

1. Create a vector `x` with the following entries:

```
3 4 1 1 2 1 4 2 1 1 5 3 1 1 1 2 4 5 5 3
```

2. Check which elements of `x` are equal to 1 (Hint: use “==” operator).
3. Create a variable (e.g vector) `y` having the logarithm of `x`.
4. Create a vector `z` containing the first five elements of `x`.
5. Create a sequence of numbers from 1 to 20 in steps of 0.2 (see help seq function).
6. Concatenate `x` and `y` into a vector called `newVec`.
7. Display all objects in the workspace and then remove `newVec` object.
8. Create a `data.frame` called `elasticband` having these two variables.

```
stretch <- c(46,54,48,50,44,42,52)
distance <- c(148,182,173,166,109,141,166)
```

9. Compute the mean of the variable `stretch` from the `elasticband` object (Hint: use “\$”).
10. Which is the class of the variable `distance`?

2 Descriptive analyses

1. Load the Orange datasets from package `datasets`.
2. Investigate the meaning of the variables (use `help` function).
3. Describe the type of variables included in this `data.frame`.
4. Get a summary of all variables.
5. Compute the correlation between `age` and `circumference`.
6. Compute the mean age of each tree (variable `Tree`).
7. The file `multicentric.txt` contains the data of a multicentric case/control study to discover risk factors of cervical cancer. Import the data in an object called `multi`.
8. How many samples and variable are in this database?
9. Describe the types of variables you have.
10. The file `multicentric.htm` contains a description of the variables. Create a table describing the number of cases and controls (variable `status`) of each country (variable `pais`) included in the study.
11. Create the same table between cases and controls and human papillomavirus (variable ‘`vph`’). Do you think that those variables are associated?

12. Summarize the duration consuming oral contraceptives (variable `durco`) in cases and controls, in each country, educative level (variable `niveledu`).