

Exercises for the course  
**“An introduction to R”**

Exercise session 6: Wednesday, March 12 2013

**Exercise 6.1:** *Supporting a hypothesis*

Apply a suitable test:

- Recall `heartbeats` from Sheet 4. Is the increase in weight of the heartbeat group compared to the control group significant? Then answer this question for every weight class separately.
- The common red spider (Bohnenspinnmilbe) is a vermin of agricultural plants to which plants react with the production of toxic substances. Do plants “remember” a former attack? Two groups of each 20 cotton plants are being infected with the mite. One group has never encountered this mite before. The second group has already survived a former attack. After a certain time, the mites on the plants are counted. You find the result in the file `mite.txt` on the web page. Is there a significant difference between the two groups? Compare the two groups visually (produce an appropriate plot) and apply a suitable test. Formulate an answer!

**Exercise 6.2:** We want to test the hypothesis that teeth enamel can be damaged due to the chlorine present in the water from public swimmingpools.

We have tested the enamel damage for 200 swimmers that train more or less than 6 hours a week.

Weekly training duration	Enamel damaged		
	Yes	No	Total
more than 6 h	29	71	100
less than 6 h	19	81	100
Total	48	152	200

a) Do this dataset support the hypothesis stated? Formulate a null-hypothesis and test it using an appropriate test. What is your conclusion?

b) Imagine we get analogous results for a double sample size.

We would have observed 200 frequent swimmers and 200 less-frequent swimmers and found enamel damage in 58 and 38 swimmers respectively.

What would then be the value of the test statistic and what would be your conclusion?

**Exercise 6.3:** The built-in data set `Puromycin` shows the reaction velocity versus substrate concentration in an enzymatic reaction involving untreated cells or cells treated with Puromycin. Load this dataset and have a look at it to see if the variables are continuous or ordinal. Apply an appropriate test to investigate the relation between the variables `conc` and `rate` for the treated and untreated cells.

Formulate your answer.

**Exercise 6.4:** The built-in data set `trees` provides the measurement of diameter (`Girth`), height and volume of black cherry trees.

The measures are expressed in inches for `Girth`, feet for `Height` and cubic feet for `Volume`. For the purpose of this exercise we will focus on `Girth` and `Height`.

Write a function `convert` that converts the values into the metric system. Your function should

take as argument the value to convert and the unit it was expressed in (inches or feet). Think about what your function should control for. Here is how it should work:

```
> convert(8.3,"inches")
```

```
[1] 0.21082
```

```
> convert(8.3,"feet")
```

```
[1] 2.52984
```

```
> convert(8.3,"foot")
```

```
Error in convert(8.3, "foot") : the unit was not "inches" or "feet"
```

Then use your function to convert the values and test if there is a significant correlation between the two variables.

Formulate your answer. Is there a difference when you use the non-converted values?