Exercices on session *Statistics with R session 1/3*

Load the package tidyverse

Admission in graduate school in UC Berkeley

We want to know if the admissions in graduate school in the Department A of UC Berkeley depends on the gender.

- The data are called *UCBAdmissions* and are available in base R. Load the data and look at their structure

- Get the data for department A (the first element of the array so you need to use the following command)

```
UCBAdmissions[, , 1]
```

- These data give the number of admitted or rejected people per gender. How is this type of table called?

- Perform a χ^2 test to see if the admission depends on the gender, and interpret it

- Check the expected frequency under the null hypothesis. Is the χ^2 test reliable?

- Get the residuals and interpret them

- Represent the residuals graphically

Weight of cabbages from different cultivars

We want to know if the weight of cabbages planted on a certain day depends on the cultivar.

- Load the dataset called *cabbages* in the package *MASS*

- Make a dataset containing only the cabbages planted at date *d21*

- Make a boxplot to graphically explore if the weight depends on the cultivar. What do you think?

- Check the normality of the distribution, using a test and graphically

- Perform a t-test

- Get the p-value, the mean estimated difference between the two groups and its confidence interval. What do you conclude?

Effect of the planting date on the content of vitamin C in cabbages

We want to know if the content of vitamin C in cabbages of the cultivar c52 depends on the planting date.

- Make a dataset containing only the cabbages of the cultivar *c52*

- Make a boxplot to explore the question graphically. What do you think?

- Perform an ANOVA

- Check the model assumptions

- Get the estimated mean of each groups

- Perform a Tukey test see which groups are different, and interpret it

- Represent the results of the ANOVA graphically

Brain and body weights in mamals

We want to know if there is a correlation between the brain and the body weights in mamals.

- Load the dataset called *mammals* in the package *MASS*

- Get the Pearson correlation coefficient

- Do a Pearson correlation test and interpret it

For next time

- Install the following packages from CRAN: *ggeffects*, *gtsummary*, *ggstats*, *lindia*, *car*, *AICcmodavg*

- Make sure the following packages are already installed: *tidyverse*, *MASS*,

- Download the *RData* file available here and store them in your raw data folder.