MAR536: Biological Statistics II

R Laboratory Exercise 7

March 01, 2023

Open a new R markdown file. Save it. (name it lastname_lab7.Rmd or something similar)
At the top of the script, add comments with your name and lab 7. Work in pairs or individually. Submit your Rmd and any other files via myCourses before lab next week.

Exercise 1

Create a function `wtlen()` with three arguments: a vector of lengths, and values `a` and `b`.
The function should return the weights of the fishes using the weight-length equation: $W = aL^b$
Use the function to calculate the weight (in g) of fish of length 100, 200, 300 cm for:

Species	a	b
Mola mola	0.0454	3.05
Regalecus glesne	0.0039	2.90

i.e. you can run your function twice, once with the paramters for *Mola* & once for *Regalecus*.

Exercise 2

Write a function `cor_vec()` which takes two vectors `X` and `Y` which:
1. Centers both vectors (subtract their respective means) so they average 0
2. Plots `X` versus `Y`
3. Returns the correlation between `X` and `Y`.

Apply your function to the following data set:
```
set.seed(66)
xx <- seq(from = 1, to = 100, by = 1)
yy <- 0.2 + xx*0.5 + rnorm(n = length(xx), mean = 0, sd = 5)
```

Add `, ...` to your function's argument list and to the call to your plot's `geom_()` command.
The `...` stands for any parameter.
Call the function with parameters that change the color & size of points.

Exercise 3

Write a function which:
1. Takes a file name and two numbers (the defaults for the two numbers should be 1 and 2),
2. Reads in a table of data (assume that the file is comma delimited),
3. Plots the columns represented by the two numbers against each other.

Hints:
Use the read_csv() function.
Use print() to check the values of intermediate results (to see if your function is working).
Use the `hake.csv` file we used previously to check your program.

Exercise 4

Write a function (or set of functions) that:
a. takes two vectors of data,

b. produces a scatterplot of one vector vs the other,

c. adds a title that is a species name,

d. performs a regression if there is enough data,

e. adds the estimated regression line to the plot, and

f. returns the coefficients of the regression.

Add an argument that determines whether the regression should be based on the logs of the input vectors.

Hint Test your functions with some dummy data. Reuse code from earlier today and from other labs.

Apply your function(s) to the `Laengelmavesi2` data.