

## BIO206 2023

### Report 1. Shared plant knowledge

#### 1) Plant knowledge file

- a) How many subjects and plants in the dataset?
- b) Calculate the total sum of knowledge **by plant** (how many people know each plant?). Are some plants more popular than others? What is the average by plant? Plot a histogram with the distribution.
- c) Calculate total knowledge **by individual** (how many plants are known by each person?) Is plant knowledge variable among individuals? What is the average plant knowledge by individual? Plot a histogram.

#### 2) Plant participants file

- a) What is the age distribution in the sample? Produce a table.
- b) Which fraction is pre-adult (5-10 or 10-15 years old)?
- c) What is the sex ratio in the sample?

#### 3) Dyads

- a) How many dyads are FF, MM, and FM?
- b) How many dyads are ‘young’ (both individuals are 15 or under) and ‘old’ (both over 60)?
- c) How many dyads born in same vs. different camp? (use *born* instead of *born\_cluster*)
- d) How many dyads were in the same vs. different camps during the interviews?

#### 4) Total knowledge score

Assuming (incorrectly) that the data points (rows) are independent, we can define ‘total score’ as the sum of shared knowledge (from 0 to number of plants) by each dyad.

Is there a difference in total score as a function of dyadic:

- a) Age? (define levels as: “young”, “old”, or “others”)
- b) Sex? (MM, MF, FF)
- c) Camp? (same vs different camp)

#### 5) Regression analysis

Now create a variable *shared knowledge* of each plant by each dyad.

Run separate logistic regressions (i.e. not controlling for pseudoreplication) to predict shared knowledge as a function of dyadic:

- a) Age
- b) Sex
- c) Camp
- d) A multiplicative model with age, sex and their interaction (age\*sex). Is the interaction significant?

Paste the optimised model outputs and interpret results a), b), c) and d)

## 6) Mixed-effects modelling

- a) Which random factor should be included in all analyses? Pick just that one, and use it in all the following analyses.

Controlling for *random intercept effects only*, run a mixed effects logistic regression model of shared knowledge on:

- b) Age
- c) Sex
- d) Camp
- e) A multiplicative model including age, sex and interaction (age\*sex)

## 7) Conclusion

Compare the results from total scores, logistic regression and mixed effects logistic regression, and discuss the effect of age, sex and camp on shared plant knowledge. Which factor(s) is/are the most important?

## 8) Code file

Attach code file you used to create dataset and run analyses here if writing the report in Word (or submit an R notebook including the code)