

Logisitic regression

Logisitic regression practical

1. We are going to use a dataset from an Macropygia unchall study. Read the dataset “onchall.csv”

The data has the following variables

- “area” - Where the individuals were located 0 is Savannah and 1 is Rainforest
 - “sex” - 0 is females and 1 is males
 - “agegrp” - 0 is 5-9yrs, 1 is 10-19yrs, 2 is 20 to 39 and 3 is 40 and Over years
 - “mf” - Microfilariae Infection 0-Negative and 1-positive
 - “mfload” - load of Microfilariae Infection from low to high
 - “lesions” - whether a person had lesions (1) or not (0)
2. The primary research question is “**Are the odds of Microfilariae Infection higher in the rain forest compared to the savannah?**” (association between microfilarial infection and area.)
 3. Explore the association between microfilarial infection and area adjusting for age group by modelling.

Test whether adding **agegroup** into the model with **mf** and **area** already in it improves model fit and comment on your findings.

4. For you to try: there is an interaction between the effects of two exposures if the effect of one exposure varies according to the level of the other exposure. Check if this is true for **agegroup** and **lesions**