

Week 3: Linear models

Remember: You can work on these in groups. Doing science socially is not cheating.

Exercise 1: Fox food

The data in the file `foxes.csv` are information about packs of urban foxes in Britain. Urban foxes forage in and defend territories. Successful groups have higher average body mass. Mass measurements are contained in the column `WEIGHT`. This is the column we want to explain, using the other columns. The meanings of the explanatory columns are given below.

`AVFOOD` Average amount of food available in the total territory.

`GSIZE` Number of adult foxes in social group.

`AREA` Kilometer area of territory.

(a) Estimate the linear model:

```
WEIGHT ~ AVFOOD
```

What is the relationship between food and body weight? For each unit increase in available food, how much heavier do foxes get?

(b) Use the command `confint()` to find the 95% confidence interval of the effect of food on body weight. How confident are you about the direction (positive or negative) and magnitude of the effect?

(c) Estimate the linear model:

```
WEIGHT ~ GSIZE
```

Again, what is the effect of group size, and what is the confidence interval?

(d) Estimate the linear model:

```
WEIGHT ~ AVFOOD + GSIZE
```

Compare the estimates and confidence intervals from this model to those from the two models above. What has happened to the estimates and why?

(e) Now, add `AREA` to the model from part (d). Why do you think `AREA` is less predictive of `WEIGHT` than the other two variables are? This is a question about biology, not just statistics.

(f) Finally, copy and fill in the table below. By “mathematical form” of each model, I mean the form that explicitly represents each parameter and variable, as well as the probability density. I do not mean the likelihood function. By “with estimates,” I mean repeating the math form of the model and replacing the parameters with their actual estimates. You can round to two decimal places.

Model	Mathematical form	With estimates
WEIGHT ~ AVFOOD		
WEIGHT ~ GSIZE		
WEIGHT ~ AVFOOD + GSIZE		
WEIGHT ~ AVFOOD + GSIZE + AREA		

Exercise 2: Predicting primates

Load the data contained in `primates.csv`. These are species identifications and body masses for some large bodied primates, living on both islands and the mainland. A rather storied hypothesis has it that island animals become smaller than mainland counterparts (unless they are very tiny, in which case they get larger, but that isn't our interest in these data, because these are all large primates).

(a) Estimate the linear model:

`Body.mass.kg ~ Genus + Type`

How important is `Genus`, compared to `Type`, for predicting body mass? There are many ways to answer this question, so think hard about it, and take your best shot. You may want to perform additional calculations or model fits, or you may find a way to use the information you already calculated.

(b) If you take a look at the data frame, you'll notice that body mass measurements are missing for 8 rows. Use the estimates from part (a) to predict the body mass measurements for these species.