

Statistical Rethinking: Notes and Selected Exercises

Somnath Sikdar

November 19, 2020

Contents

2	Small Worlds and Large Worlds	2
---	-------------------------------	---

Chapter 2

Small Worlds and Large Worlds

Exercise 2.1. Suppose there are two species of panda bear. Both are equally common in the wild and live in the same places. They look exactly alike and eat the same food, and there is yet no genetic assay capable of telling them apart. They differ however in their family sizes. Species A gives birth to twins 10% of the time, otherwise birthing a single infant. Species B births twins 20% of the time, otherwise birthing singleton infants. Assume these numbers are known with certainty, from many years of field research. Now suppose you are managing a captive panda breeding program. You have a new female panda of unknown species, and she has just given birth to twins. What is the probability that her next birth will also be twins?

Solution. We have to estimate $\Pr\{\text{twins again}|\text{twins before}\}$. We may write this conditional probability as follows:

$$\begin{aligned}\Pr\{\text{twins again}|\text{twins before}\} &= \Pr\{\text{twins again}|A, \text{twins before}\} \cdot \Pr\{A|\text{twins before}\} + \\ &\quad \Pr\{\text{twins again}|B, \text{twins before}\} \cdot \Pr\{B|\text{twins before}\} \\ &= \Pr\{\text{twins again}|A\} \cdot \Pr\{A|\text{twins before}\} + \\ &\quad \Pr\{\text{twins again}|B\} \cdot \Pr\{B|\text{twins before}\}\end{aligned}\tag{2.1}$$

The last equality follows, since when we condition on any given species, the probability of having twins again is independent of whether twins were born before. Strictly speaking, this is also an assumption but a reasonable one. As such, we do not use the “before” and “again” qualifiers. To evaluate the last expression, we calculate $\Pr\{\cdot|\text{twins}\}$ for both species A and B . Using Bayes’ Theorem, we obtain:

$$\begin{aligned}\Pr\{A|\text{twins}\} &= \frac{\Pr\{A|\text{twins}\} \cdot \Pr\{A\}}{\Pr\{A|\text{twins}\} \cdot \Pr\{A\} + \Pr\{B|\text{twins}\} \cdot \Pr\{B\}} \\ &= \frac{0.10 \times 0.5}{0.5 \times (0.1 + 0.20)} \\ &= \frac{1}{3}.\end{aligned}$$

A similar calculation yields $\Pr\{B|\text{twins}\} = \frac{2}{3}$. The final expression in (2.1) evaluates to

$$0.10 \times \frac{1}{3} + 0.20 \times \frac{2}{3} = \frac{1}{6} = 0.17.$$

This probability lies between 0.10 and 0.20 as expected and lies closer to 0.20 than 0.10. This is also clear since if the panda has had twins before, it is more likely to be of species B than species A . Hence it is more likely that she will birth twins again.