Statistical Rethinking

Chapter 2

Here are my solutions to the problems in sections 'medium' and 'hard'.

Medium

2M3. Suppose there are two globes, one for Earth and one for Mars. The Earth globe is 70% covered in water. The Mars globe is 100% land. Further suppose that one of these globes—you don't know which—was tossed in the air and produced a "land" observation. Assume that each globe was equally likely to be tossed. Show that the posterior probability that the globe was the Earth, conditional on seeing "land" (Pr(Earth|land)), is 0.23.

```
* P(W|E) = 0.7
* P(L|E) = 1 - 0.7= 0.3
* P(L|M) = 1
```

$$P(E|L) = \frac{P(L|E) * P(E)}{P(L)} = \frac{0.3 * 0.5}{0.3 * 0.5 + 0.5 * 1} = 0.23$$

Hard

2H1. 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?

- P(A) = 0.5
- P(B) = 0.5
- P(T|A) = 0.1
- P(T|B) = 0.2
- P(T) = $0.1 * 0.5 + 0.2 * 0.5 = 0.15$

The probability that the panda was of type B, given the bear gave birth to twins :

$$P(B|T) = \frac{P(T|B) * P(B)}{P(T)} = \frac{0.2 * 0.5}{0.15} = 0.66$$

The probability that the panda was of type A, given the bear gave birth to twins :

$$P(A|T) = \frac{P(T|A) * P(A)}{P(T)} = \frac{0.1 * 0.5}{0.15} = 0.33$$

The probability of having twins for a second time is the sum of the probability of a giving birth to twins again plus the probability of b giving birth again. This should be multiplied by the probabilities of the bear being A or B after having twins. Here are the calculations:

- $P(TT) = 0.66 * 0.2 + 0.33 * 0.1 = 0.166$

2H3. Continuing on from the previous problem, suppose the same panda mother has a second birth and that it is not twins, but a singleton infant. Compute the posterior probability that this panda is species A.

The probability of having a singleton after having twins is 1 minus the probability of having twins twice, which we calculated above. So:

- $P(TS) = 1 - P(TT) = 1 - 0.166 = 0.84$

$$P(A|TS) = \frac{P(TS|A) * P(A)}{P(TS)} = \frac{0.9 * 0.33}{0.84} = 0.36$$

2H4. A common boast of Bayesian statisticians is that Bayesian inference makes it easy to use all of the data, even if the data are of different types. So suppose now that a veterinarian comes along who has a new genetic test that she claims can identify the species of our mother panda. But the test, like all tests, is imperfect. This is the information you have about the test: * The probability it correctly identifies a species A panda is 0.8. * The probability it correctly identifies a species B panda is 0.65. The vet administers the test to your panda and tells you that the test is positive for species A. First ignore your previous information from the births and compute the posterior probability that your panda is species A.

- P(+|A) = 0.8
- P(+|B) = 0.65
- P(+) $= 0.8 * 0.5 + (1 - 0.65) * 0.5 = 0.575$

$$P(A|+) = \frac{P(+|A) * P(A)}{P(+)} = \frac{0.8 * 0.5}{0.575} = 0.69$$

Then redo your calculation, now using the birth data as well.

- P(+) $= 0.8 * 0.36 + (1 - 0.65) * 0.64$= 0.504

$$P(TSA|+) = \frac{P(+|TSA) * P(TSA)}{P(+)} = \frac{0.36 * 0.8}{0.504} = 0.57$$