1 Easy

1.1

Which of the expressions below correspond to the statement: *the probability of rain on Monday*?

1. $Pr(rain)$
2. $Pr(rain|Monday)$
3. $Pr(Monday|rain)$
4. $Pr(rain, Monday)/Pr(Monday)$

Understanding the question to mean "*the probability of rain given that it is Monday,*" not "*probability of Monday and rain,*" **1.** $Pr(rain|Monday)$ corresponds to the statement as well as **4.** $Pr(rain, Monday)/Pr(Monday)$.

1.2

Which of the following statements corresponds to the expression: $Pr(Monday|rain)$?

1. The probability of rain on Monday.
2. The probability of rain, given that it is Monday.
3. The probability that it is Monday, given that it is raining.
4. The probability that it is Monday and that it is raining.

3. is the only option that corresponds to the expression.

1.3

Which of the expressions below correspond to the statement: *the probability that it is Monday, given that it is raining*?

1. $Pr(Monday|rain)$
2. $Pr(rain|Monday)$
3. $Pr(rain|Monday)Pr(Monday)$
4. $Pr(rain|Monday)Pr(Monday)/Pr(rain)$
5. $Pr(Monday|rain)Pr(rain)/Pr(Monday)$

1. corresponds to the statement directly, **4.** is equivalent by Bayes theorem.

1.4

The Bayesian statistician Bruno de Finetti (1906–1985) began his book on probability theory with the declaration: "PROBABILITY DOES NOT EXIST." The capitals appeared in the original, so I imagine de Finetti wanted us to shout this statement. What he meant is that probability is a device for describing uncertainty from the perspective of an observer with limited knowledge; it has no objective reality. Discuss the globe tossing example from the chapter, in light of this statement. What does it mean to say "the probability of water is 0.7"?

The "probability" is a proxy for the actual ratio of water on the globe given our sampling technique; while the globe can be shown to actually be 70% covered by water, successive tosses are either water or land, not somewhere in between.

2 Medium

2.1

m1.r

2.2

m2.r

2.3

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.

$$Pr(Earth|land) = \frac{Pr(land|Earth) * Pr(Earth)}{Pr(land)} \tag{1}$$

$$Pr(Earth|land) = \frac{0.3 * 0.5}{0.5 * (1 + 0.3)} \tag{2}$$

$$Pr(Earth|land) \simeq 0.23 \tag{3}$$

2.4

Suppose you have a deck with only three cards. Each card has two sides, and each side is either black or white. One card has two black sides. The second card has one black and one white side. The third card has two white sides. Now suppose all three cards are placed in a bag and shuffled. Someone reaches into the bag and pulls out a card and places it flat on a table. A black side is shown facing up, but you don't know the color of the side facing down. Show that the probability that the other side is also black is 2/3. Use the counting method (Section 2 of the chapter) to approach this problem. This means counting up the ways that each card could produce the observed data (a black side facing up on the table).

Call the cards BB, BW, and WW. If the card is WW, we have 0 ways of having one side be black. If the card is BW, we have 1 way of having one side be black. If the card is BB, we have 2 ways of having one side be black. So, we have 3 total ways of having the shown side be black, 2 of which correspond to BB, implying a 2/3 probability of the card being BB.

2.5

Now suppose there are four cards: B/B, B/W, W/W, and another B/B. Again suppose a card is drawn from the bag and a black side appears face up. Again calculate the probability that the other side is black.

The number of ways for BB_1 and BB_2 to be black are both 2, for BW there is only one way. So, we have 5 total ways of drawing a black, 4 of which correspond to BB type cards, implying a 4/5 probability of the other side being black.

2.6

Imagine that black ink is heavy, and so cards with black sides are heavier than cards with white sides. As a result, it's less likely that a card with black sides is pulled from the bag. So again assume there are three cards: B/B, B/W, and W/W. After experimenting a number of times, you conclude that for every way to pull the B/B card from the bag, there are 2 ways to pull the B/W card and 3 ways to pull the W/W card. Again suppose that a card is pulled and a black side appears face up. Show that the probability the other side is black is now 0.5. Use the counting method, as before.

Have $1 * n_B B + 2 * n_B W = 1 * 2 + 2 * 1 = 4$ total ways to draw a card with a black side face up, of which only $n_B B = 2$ correspond to a black side underneath, implying a $2/4 = 0.5$ probality.

2.7

Assume again the original card problem, with a single card showing a black side face up. Before looking at the other side, we draw another card from the bag and lay it face up on the table. The face that is shown on the new card is white. Show that the probability that the first card, the one showing a black side, has black on its other side is now 0.75. Use the counting method, if you can. Hint: Treat this like the sequence of globe tosses, counting all the ways to see each observation, for each possible first card.

On the first draw have 2 ways of getting BB, 1 way of getting BW. On the second draw, have 1 way of getting BW, 2 ways of getting WW. So, have the following numbers of ways of getting the sequences:

$$BB, BW \rightarrow 2 * 1 = 2 \quad (4)$$
$$BB, WW \rightarrow 2 * 2 = 4 \quad (5)$$
$$BW, BW \rightarrow 1 * 0 = 0 \quad (6)$$
$$BW, WW \rightarrow 1 * 2 = 2 \quad (7)$$
$$(8)$$

So, 8 ways of drawing the B, W sequence in total, 6 of which correspond to an initial draw of BB, implying a 0.75 probability of having a black side underneath on the first card.

3 Hard

3.1

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?

Denote A, B as event that panda is species A, B, T as twins event.

$$Pr(A|T) = \frac{Pr(T|A) * Pr(A)}{Pr(T)} \tag{9}$$

$$Pr(A|T) = \frac{0.1 * 0.5}{0.5 * (0.1 + 0.2)} \tag{10}$$

$$Pr(A|T) = 1/3, Pr(B|T) = 2/3 \tag{11}$$

So, expected probability of next litter being twins is

$$1/3 * Pr(T|A) + 2/3 * Pr(T|B) = \tag{12}$$

$$1/3 * 0.1 + 2/3 * 0.2 \simeq 16.7\% \tag{13}$$

3.2

Recall all the facts from the problem above. Now compute the probability that the panda we have is from species A, assuming we have observed only the first birth and that it was twins.

$Pr(A|T) = 1/3$ as stated before.

3.3

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.

After having the twins, we update our priors $P(A) = 1/3$, $Pr(B) = 2/3$. Find the new $Pr(A|T^c)$:

$$Pr(A|T^c) = \frac{Pr(T^c|A) * Pr(A)}{Pr(T^c)} \tag{14}$$

$$Pr(A|T^c) = \frac{0.9 * 1/3}{1/3 * 0.9 + 2/3 * 0.8} \tag{15}$$

$$Pr(A|T^c) = 36\% \tag{16}$$

3.4

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

Denote the test showing species A as α, species B as β. We know that $Pr(\alpha|A) = 0.8$, $Pr(\beta|B) = 0.65$. Let's find $Pr(A|\alpha)$ with the initial priors $Pr(A) = Pr(B) = 0.5$:

$$Pr(A|\alpha) = \frac{Pr(\alpha|A) * Pr(A)}{Pr(\alpha)} \tag{17}$$

$$Pr(A|\alpha) = \frac{0.8 * 0.5}{0.5 * 0.8 + 0.5 * 0.35} \tag{18}$$

$$Pr(A|\alpha) \simeq 70\% \tag{19}$$

Now let's recalculate $Pr(A|\alpha)$ with the new priors $Pr(A) = 0.36$, $Pr(B) = 0.64$.

$$Pr(A|\alpha) = \frac{Pr(\alpha|A) * Pr(A)}{Pr(\alpha)} \tag{20}$$

$$Pr(A|\alpha) = \frac{0.8 * 0.36}{0.36 * 0.8 + 0.64 * 0.35} \tag{21}$$

$$Pr(A|\alpha) \simeq 56\% \tag{22}$$