hw2

Christopher Huong

2024-01-25

Q. 2E1. Which of the expressions below correspond to the statement: the probability of rain on Monday?

(1) P(rain|Monday)

Q. 2E2. Which of the following statements corresponds to the expression: P(Monday|rain)?

(3) The probability that it is Monday, given it is raining

Q. 2E3. Which of the expressions below correspond to the statement: the probability that it is Monday, given that it is raining?

(1) P(Monday|rain)

Q. 2E4. …. 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 observer's strength of belief that the next toss will result in water is 70%.

Q. 2M1. Recall the globe tossing model from the chapter. Compute and plot the grid approximate posterior distribution for each of the following sets of observations. In each case, assume a uniform prior for p.

1. W, W, W

```
n=11 #10 sided globe
#make list of parameter values on the grid
p_grid <- seq(from=0, to=1, length.out=n)
#uniform prior at each parameter value on grid
prior <- rep(1,n)
#compute likelihood at each parameter value
likelihood <- dbinom(x=3, size=3, prob=p_grid)
unstd.posterior <- prior * likelihood
posterior <- unstd.posterior / sum(unstd.posterior)
plot( p_grid , posterior , type="b" )
```

2. W, W, W, L

```
likelihood <- dbinom(x=3, size=4, prob=p_grid)
unstd.posterior <- prior * likelihood
posterior <- unstd.posterior / sum(unstd.posterior)
plot( p_grid , posterior , type="b" )
```
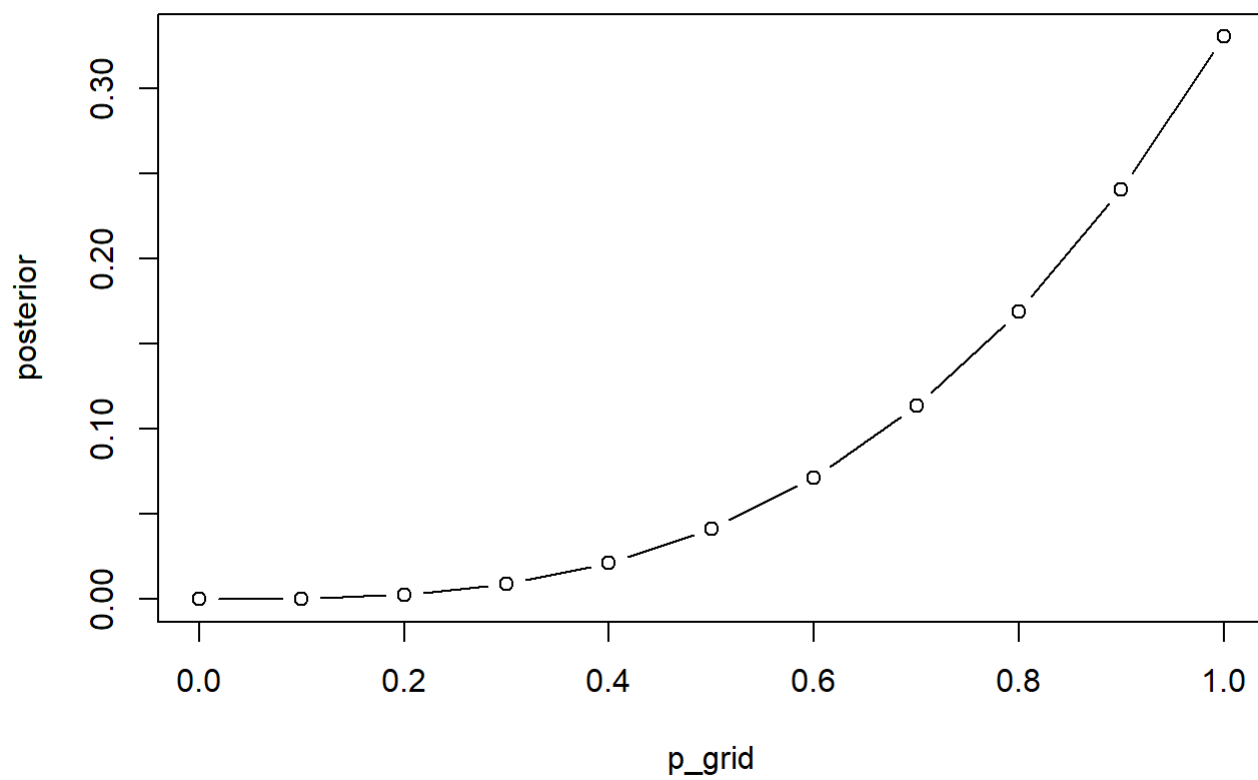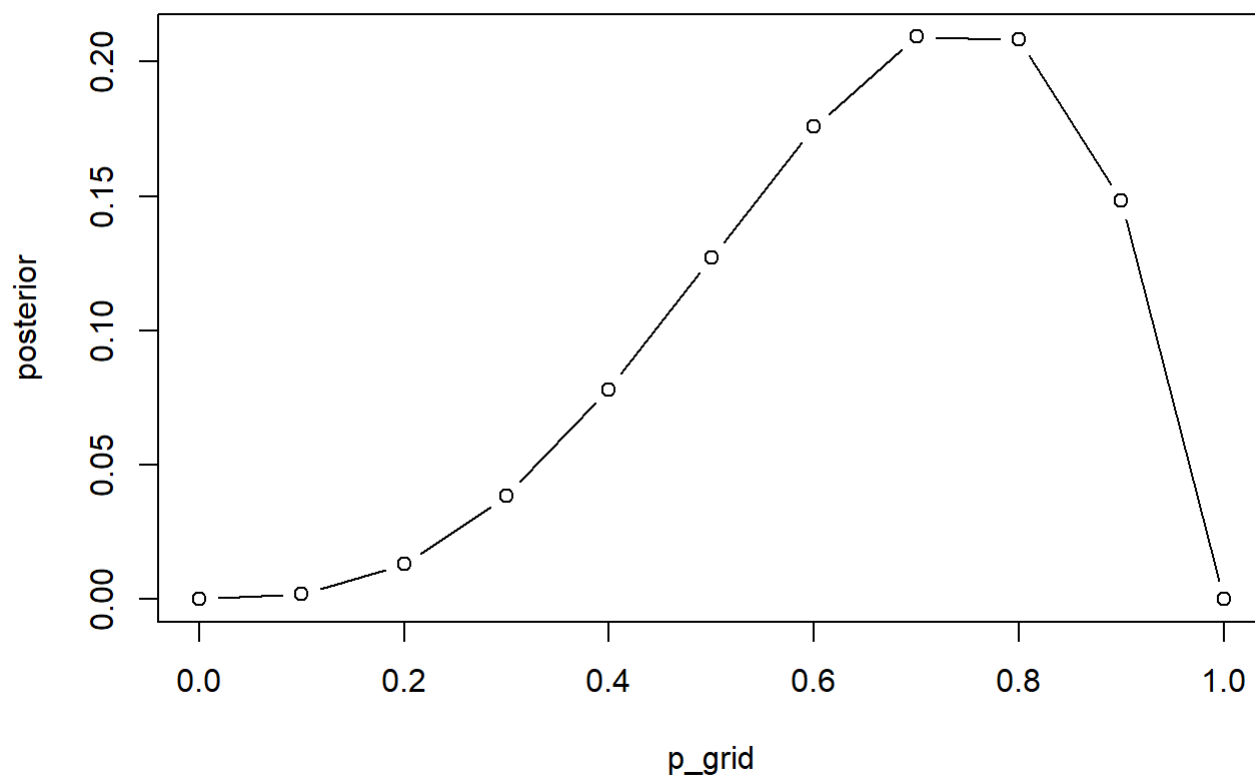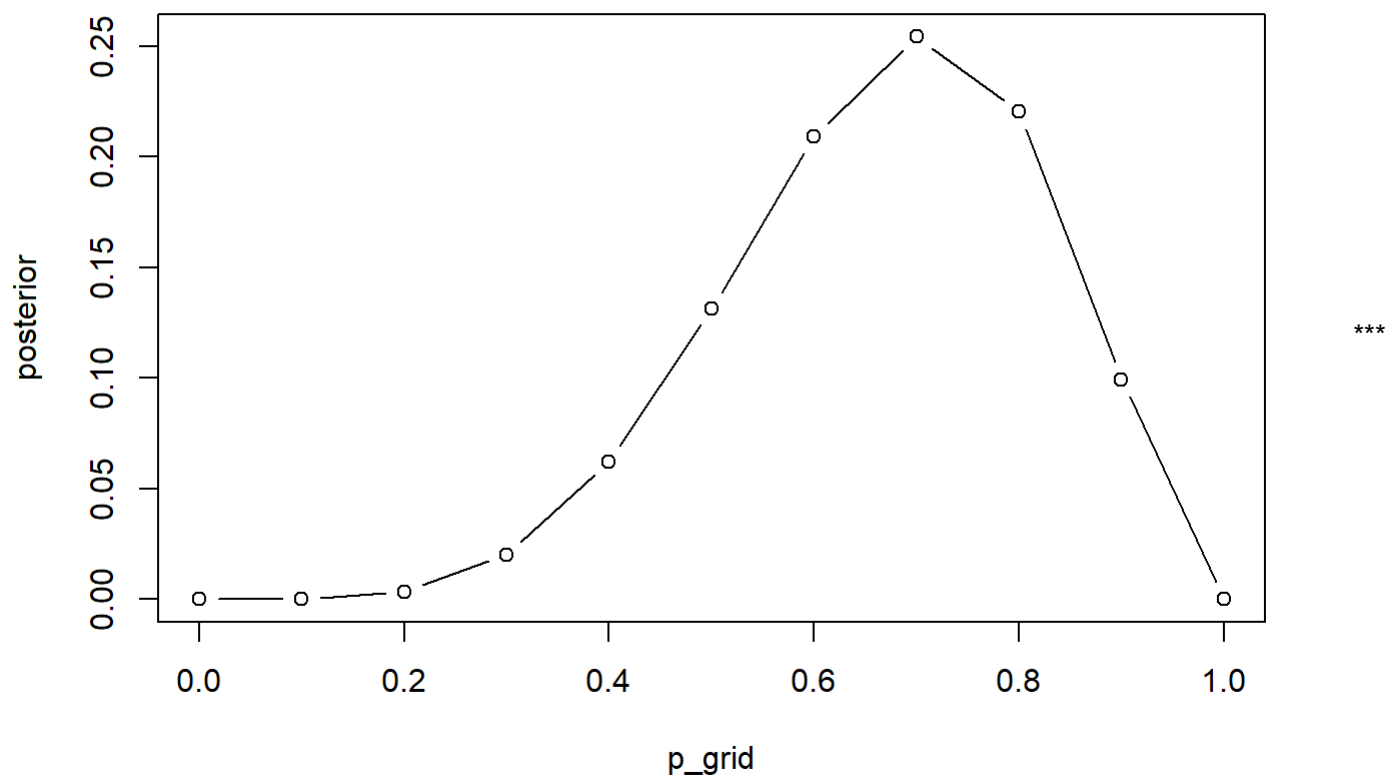
3. L, W, W, L, W, W, W

```
likelihood <- dbinom(x=5, size=7, prob=p_grid)
unstd.posterior <- prior * likelihood
posterior <- unstd.posterior / sum(unstd.posterior)
plot( p_grid , posterior , type="b" )
```

Q. 2M2. Now assume a prior for p that is equal to zero when p < 0.5 and is a positive constant when p ≥ 0.5. Again compute and plot the grid approximate posterior distribution for each of the sets of observations in the problem just above.

1. W, W, W

```
prior <- ifelse(p_grid < 0.5, 0, 1)

likelihood <- dbinom(x=3, size=3, prob=p_grid)
unstd.posterior <- prior * likelihood
posterior <- unstd.posterior / sum(unstd.posterior)
plot( p_grid , posterior , type="b" )
```
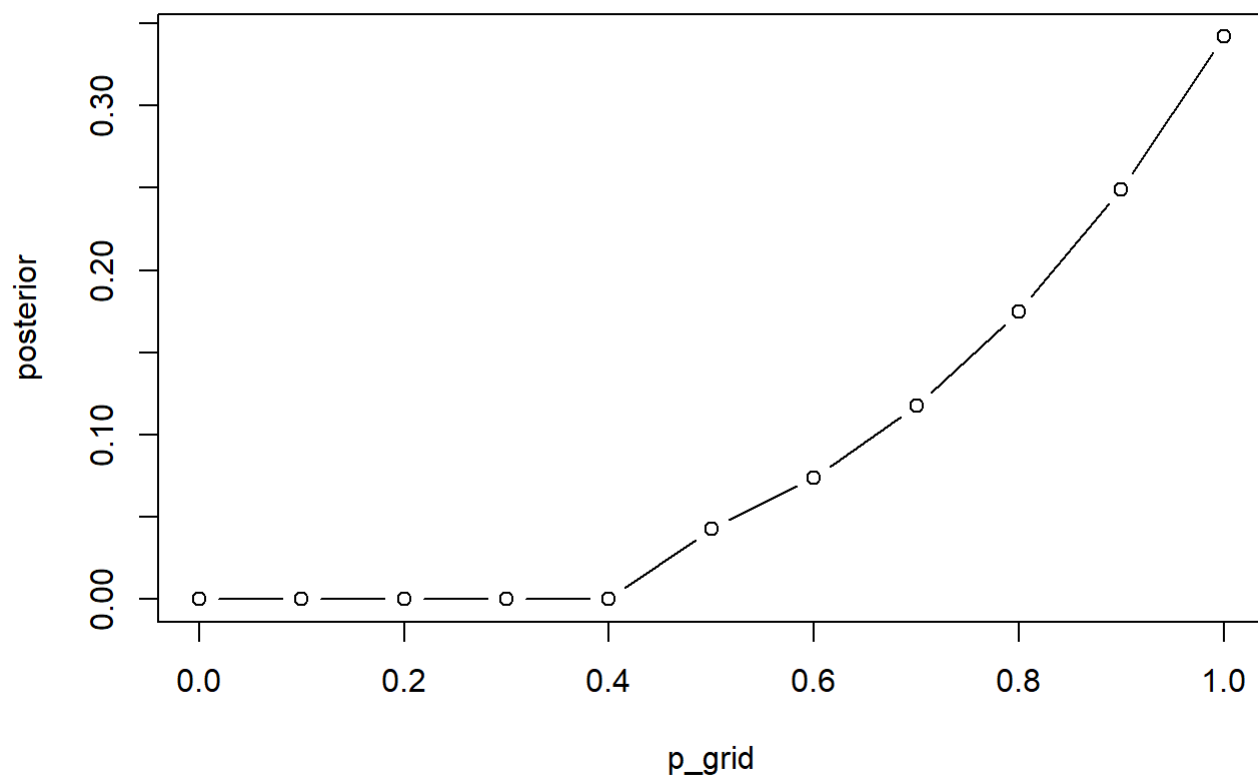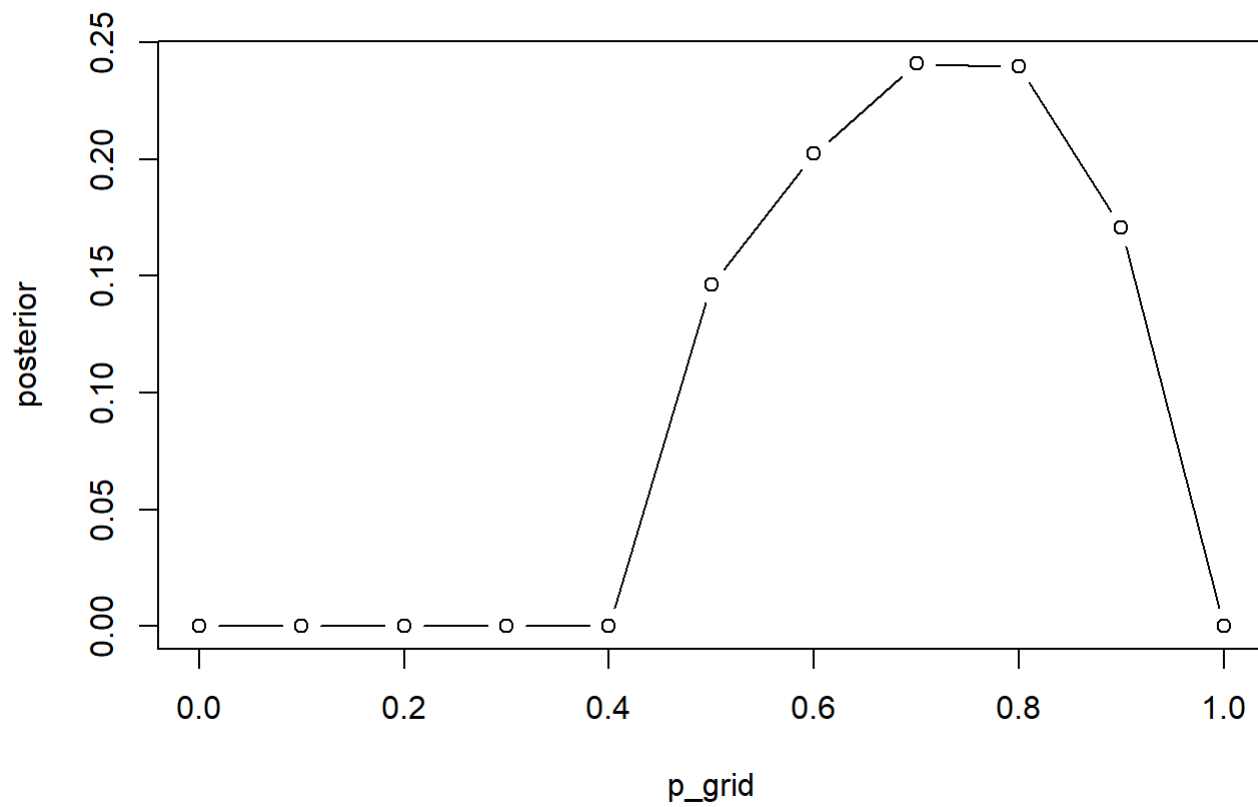
2. W, W, W, L

```
likelihood <- dbinom(x=3, size=4, prob=p_grid)
unstd.posterior <- prior * likelihood
posterior <- unstd.posterior / sum(unstd.posterior)
plot( p_grid , posterior , type="b" )
```

3. L, W, W, L, W, W, W

```
likelihood <- dbinom(x=5, size=7, prob=p_grid)
unstd.posterior <- prior * likelihood
posterior <- unstd.posterior / sum(unstd.posterior)
plot( p_grid , posterior , type="b" )
```
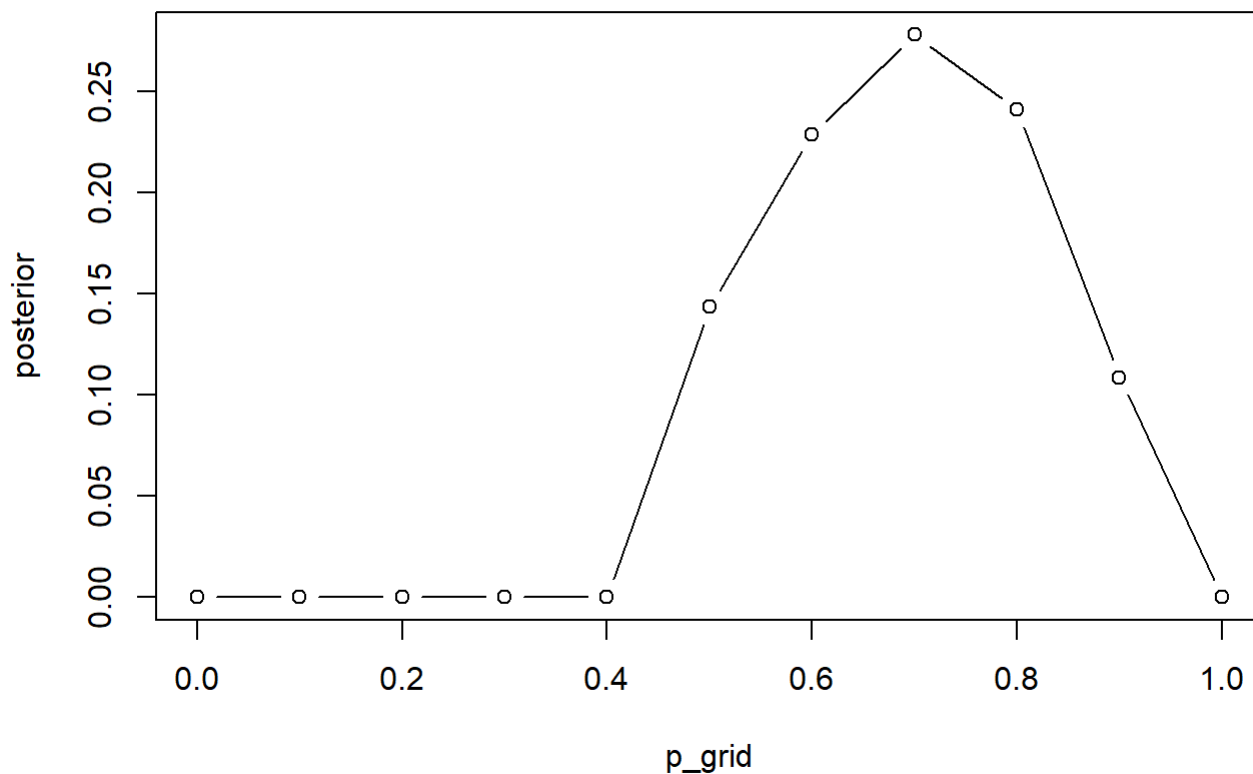
Q. 2M3. Suppose there are two globes, one for Earth and one for Mars. The Earth globe is 70% covered in water. The Mars globe is 100% land. Further suppose that one of these globes—you don't know which— was tossed in the air and produced a "land" observation. Assume that each globe was equally likely to be tossed. Show that the posterior probability that the globe was the Earth, conditional on seeing "land" (Pr(Earth|land)), is 0.23.

Given: P(land|Earth) = 0.3
P(land|Mars) = 1.0
P(Earth) = 0.5
P(Mars) = 0.5
P(land) = P(land|Earth) * P(Earth) + P(land|Mars) * P(Mars)

Bayes theorem: P(A|B) = P(B|A)P(A) / P(B)
P(Earth|land) = (P(land|Earth) P(Earth)) / P(land)

```
p_land = 0.3 * 0.5 + 1.0 * 0.5
p_land
```

```
## [1] 0.65
```

```
p_earth_given_land = (0.3 * 0.5) / 0.65
p_earth_given_land
```

```
## [1] 0.2307692
```

Q. 2M4. 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).

Given:
Card 1 = 2 black 0 white = 2 ways to produce the observed data
Card 2 = 1 black 1 white = 1 way to produce the observed data
Card 3 = 0 black 2 white = 0 ways to produce the observed data

Total 3 ways to produce the observed data
P(Card1|Black) = 2/3
P(Card2|Black) = 1/3
P(Card3|Black) = 0/3

Only Card1 will have the other side black, thus that probability is 2/3

Q. 2M5. 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.

Given:
Card 1 = 2 black 0 white = 2 ways to produce the observed data
Card 2 = 1 black 1 white = 1 way to produce the observed data
Card 3 = 0 black 2 white = 0 ways to produce the observed data
Card 4 = 1 black 1 white = 2 ways to produce the observed data

Total 5 ways to produce the observed data

P(Card1|Black) = 2/5
P(Card2|Black) = 1/5
P(Card3|Black) = 0/5
P(Card4|Black) = 2/5

Cards 1 and 4 will have the other side also black, thus that probability is 4/5

Q. 2M6. 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

Given:
Card 1 = 2 black 0 white = 2 ways to produce observed data
Card 2 = 1 black 1 white *2* = *2 ways to produce observed data*
Card 3 = 0 black 2 white 3 = 0 ways to produce observed data

Of the 4 ways to produce observed data (black side up), 2 of those also have a black side down. 2/4 = 0.5

Q. 2M7. 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.

Given:
Card 1 = 2 black 0 white
Card 2 = 1 black 1 white
Card 3 = 0 black 2 white

Possible draws to produce the observed data (black, then white)
Card1 (B1), Card2 (W)
Card1 (B2), Card2 (W)
Card1 (B1), Card3 (W1)
Card1 (B1), Card3 (W2)
Card1 (B2), Card3 (W1)
Card1 (B2), Card3 (W2)
Card2 (B), Card3 (W1)
Card2 (B), Card3 (W2)

Of the 8 ways to produce the data, 6 of them include Card 1 as the card, which would have the other side of the first card be black as well.

Thus that probability is 6/8 = 0.75

Q. 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?

Given:
P(PandaA) = 0.5
P(PandaB) = 0.5
P(Twins|PandaA) = 0.10
P(Twins|PandaB) = 0.20
P(Twins) = 0.15

Bayes theorem: P(A|B) = P(B|A)*P(A) / P(B)

Compute posterior probability of the panda being species A and B given she just birthed twins

P(PandaA|Twins) = P(Twins|PandaA) * P(PandaA) / P(Twins)
= (0.10 * 0.5) / 0.15 = 0.333

P(PandaB|Twins) = P(Twins|PandaB) * P(PandaB) / P(Twins)
= (0.20 * 0.5) / 0.15 = 0.667

Compute updated probabilities of twins for each species, given twins were just observed

P(Twins2) = P(Twins|PandaA) * P(PandaA) + P(Twins|Panda B) * P(PandaB)

= 0.10 * 0.333 + 0.20 * 0.667

Q. 2H2. 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

A) P(PandaA|Twins) = P(Twins|PandaA) * P(PandaA) / P(Twins)

= (0.10 * 0.5) / 0.15 = 0.333

Q. 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

P('Twins) = P('Twins|PandaA) * P(PandaA) + P('Twins|PandaB) * P(PandaB)
= (1-.10) * (0.333) + (1-0.20) * (0.667) = 0.833

$$P(PandaA|'Twins) = P('Twins|PandaA) * P(PandaA) / P('Twins)$$

$$= (0.90 * 0.333) / 0.833 = 0.360$$

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

Given:
P(TestA|PandaA) = 0.80
P(TestA|PandaB) = 1-0.65 = 0.35
P(TestA) = 0.80 * 0.50 + 0.35 * 0.50 = 0.575

P(PandaA|TestA) = P(TestA|PandaA) * P(PandaA) / P(TestA)
= 0.80 * 0.50 / 0.575 = 0.696