

Statistical Rethinking: Chapter 2 - Small Worlds and Large Worlds

Chris Hayduk

December 28, 2018

Problem 2E1.

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

1. $\text{Pr}(\text{rain})$
2. $\text{Pr}(\text{rain}|\text{Monday})$
3. $\text{Pr}(\text{Monday}|\text{rain})$
4. $\text{Pr}(\text{rain, Monday})/\text{Pr}(\text{Monday})$

$\text{Pr}(\text{rain}|\text{Monday})$ and $\text{Pr}(\text{rain, Monday})/\text{Pr}(\text{Monday})$ both correspond to the statement. In English, they represent the probability that, given that the day is Monday, it will rain.

Since $\text{Pr}(\text{rain, Monday}) = \text{Pr}(\text{rain}|\text{Monday}) * \text{Pr}(\text{Monday})$, $\text{Pr}(\text{rain}|\text{Monday})$ and $\text{Pr}(\text{rain, Monday})/\text{Pr}(\text{Monday})$ are equivalent.

Problem 2E2.

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

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.

$\text{Pr}(\text{Monday}|\text{rain})$ corresponds to the statement *The probability that it is Monday, given that it is raining.*

Problem 2E3.

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

1. $\text{Pr}(\text{Monday}|\text{rain})$

2. $\Pr(\text{rain}|\text{Monday})$
3. $\Pr(\text{rain}|\text{Monday}) * \Pr(\text{Monday})$
4. $\Pr(\text{rain}|\text{Monday}) * \Pr(\text{Monday}) / \Pr(\text{rain})$
5. $\Pr(\text{Monday}|\text{rain}) * \Pr(\text{rain}) / \Pr(\text{Monday})$

$\Pr(\text{Monday}|\text{rain})$ and $\Pr(\text{rain}|\text{Monday}) * \Pr(\text{Monday}) / \Pr(\text{rain})$ correspond to the statement.

Problem 2E4.

The Bayesian statistician Bruno de Finetti (1906-1985) began his book on probability theory with the declaration: "PROBABILITY DOES NOT EXIST". 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"?

In light of this statement, "the probability of water is 0.7" means that, given a random point on Earth, there is a 70% that it is water. The source of the uncertainty and limited knowledge is that the point is random - the observer cannot know whether there is land or water at the given point.

Problem 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
2. W, W, W, L
3. L, W, W, L, W, W, W

```
#Grid of probability values
p_grid <- seq(from = 0, to = 1, length.out = 100)

#Uniform prior - every value is 1
prior <- rep(1, length(p_grid))

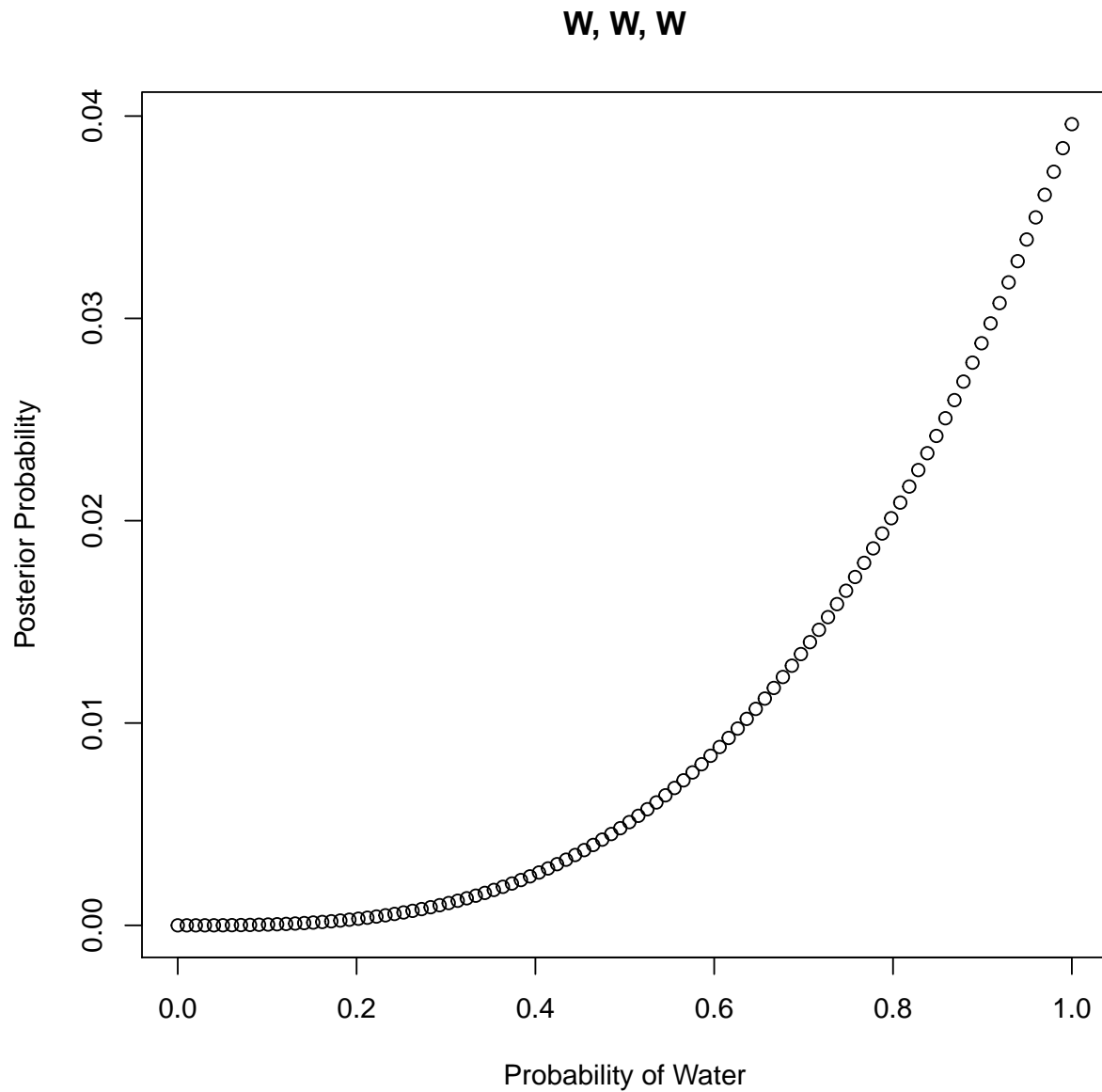
#W, W, W

likelihood <- dbinom(3, size = 3, prob = p_grid)

unstd.posterior <- likelihood * prior
```

```
posterior <- unstd.posterior/sum(unstd.posterior)

plot(p_grid, posterior, type = "b", xlab = "Probability of Water",
     ylab = "Posterior Probability", main = "W, W, W")
```



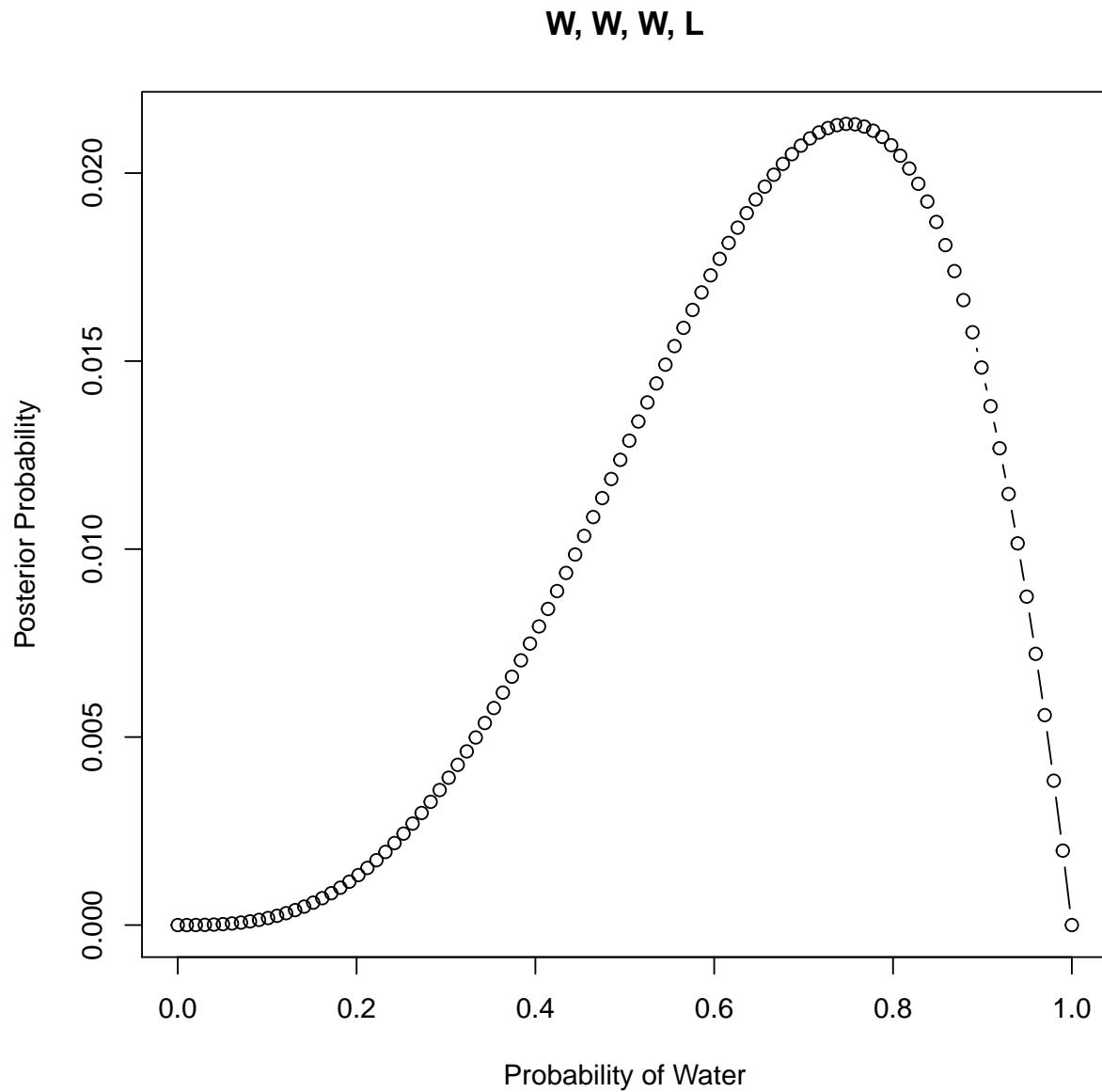
```
#W, W, W, L

likelihood <- dbinom(3, size = 4, prob = p_grid)

unstd.posterior <- likelihood * prior
```

```
posterior <- unstd.posterior/sum(unstd.posterior)

plot(p_grid, posterior, type = "b", xlab = "Probability of Water",
     ylab = "Posterior Probability", main = "W, W, W, L")
```



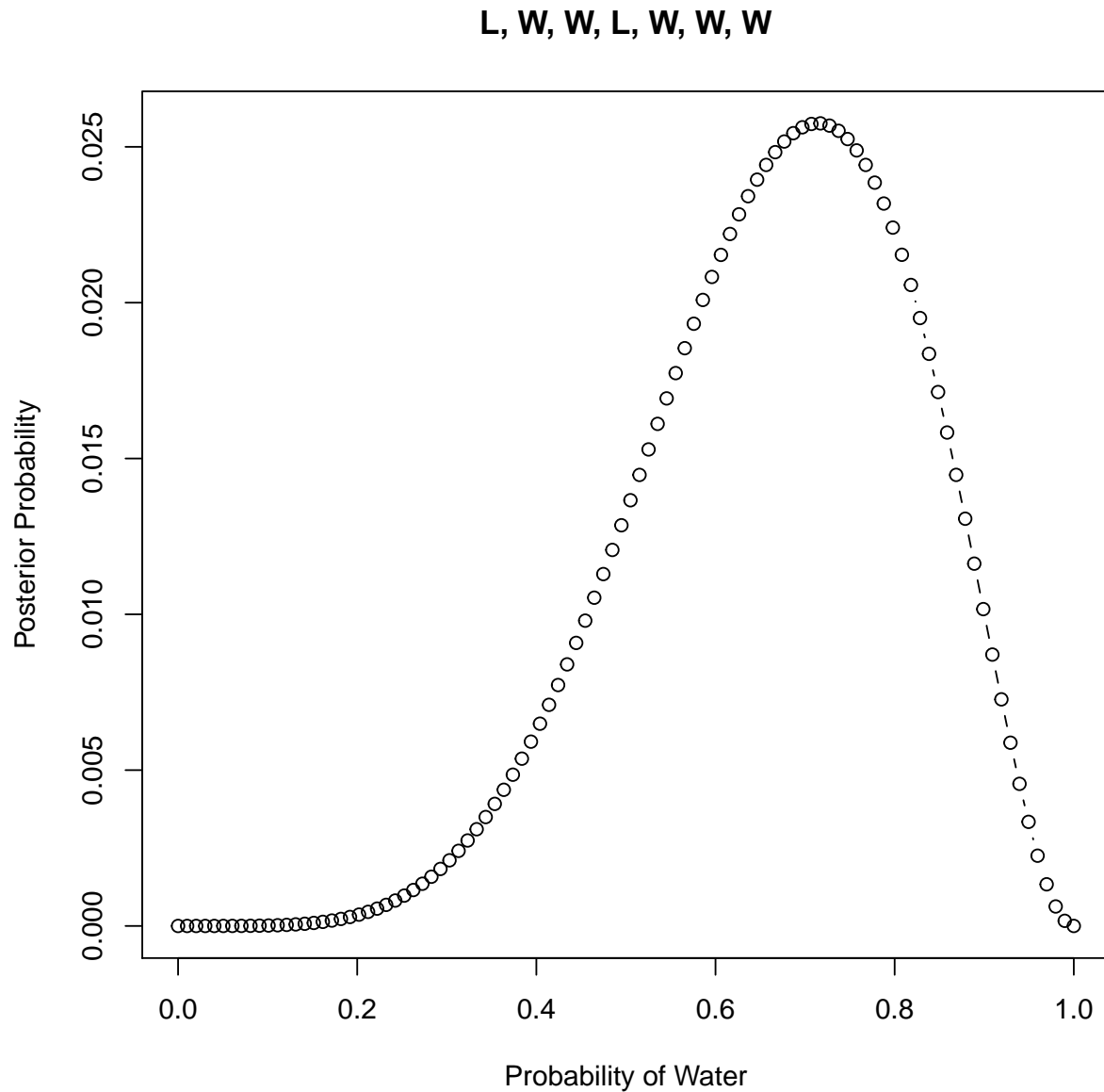
```
#L, W, W, L, W, W, W

likelihood <- dbinom(5, size = 7, prob = p_grid)

unstd.posterior <- likelihood * prior
```

```
posterior <- unstd.posterior/sum(unstd.posterior)

plot(p_grid, posterior, type = "b", xlab = "Probability of Water",
     ylab = "Posterior Probability", main = "L, W, W, L, W, W, W")
```



Problem 2M2.

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

```

#Grid of probability values
p_grid <- seq(from = 0, to = 1, length.out = 100)

#P_grid values below 0.5 receive a 0 prior,
#any p_grid value 0.5 or above is assigned a prior of 1
prior <- ifelse(p_grid < 0.5, 0, 1)

#W, W, W

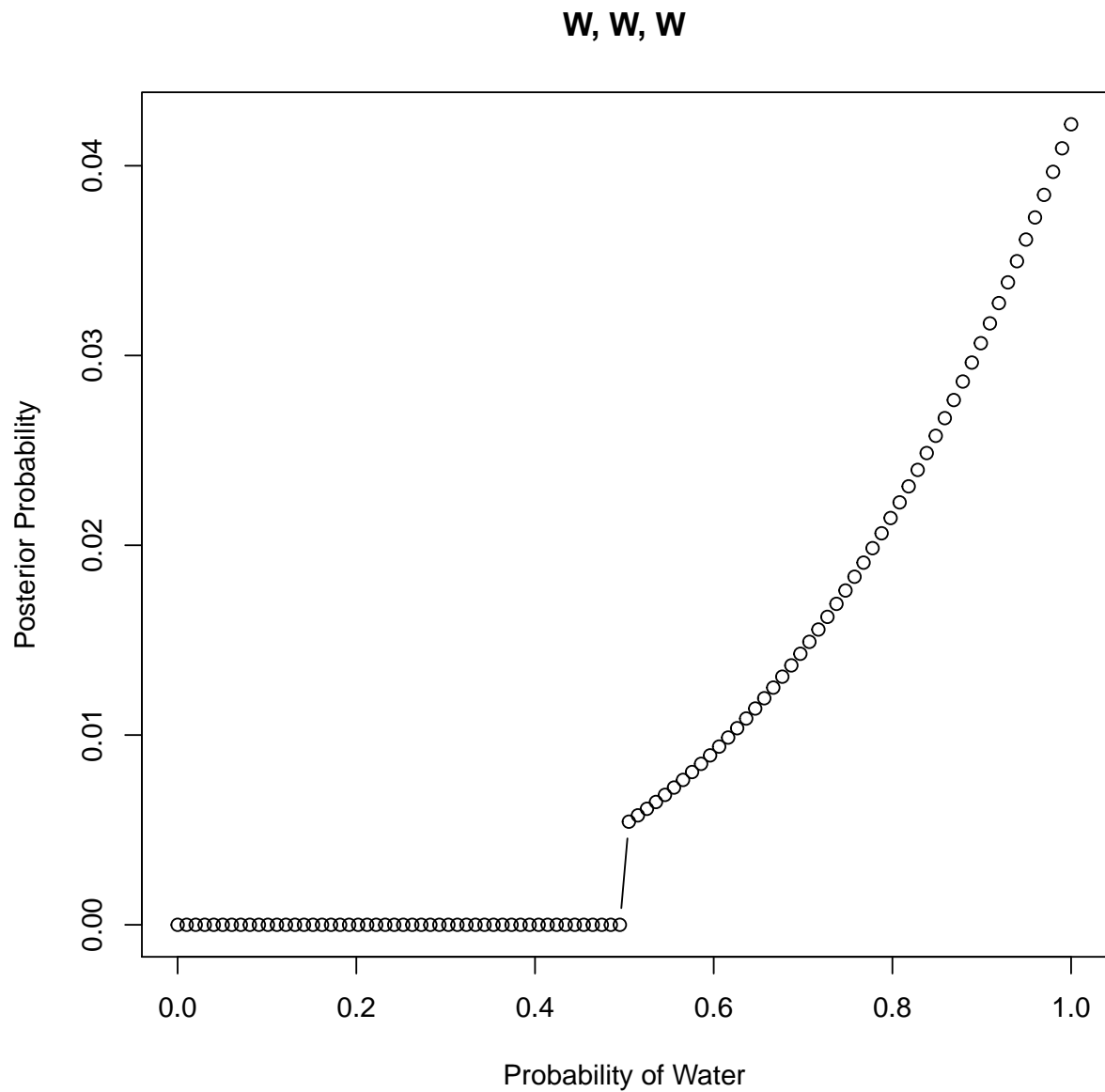
likelihood <- dbinom(3, size = 3, prob = p_grid)

unstd.posterior <- likelihood * prior

posterior <- unstd.posterior/sum(unstd.posterior)

plot(p_grid, posterior, type = "b", xlab = "Probability of Water",
      ylab = "Posterior Probability", main = "W, W, W")

```



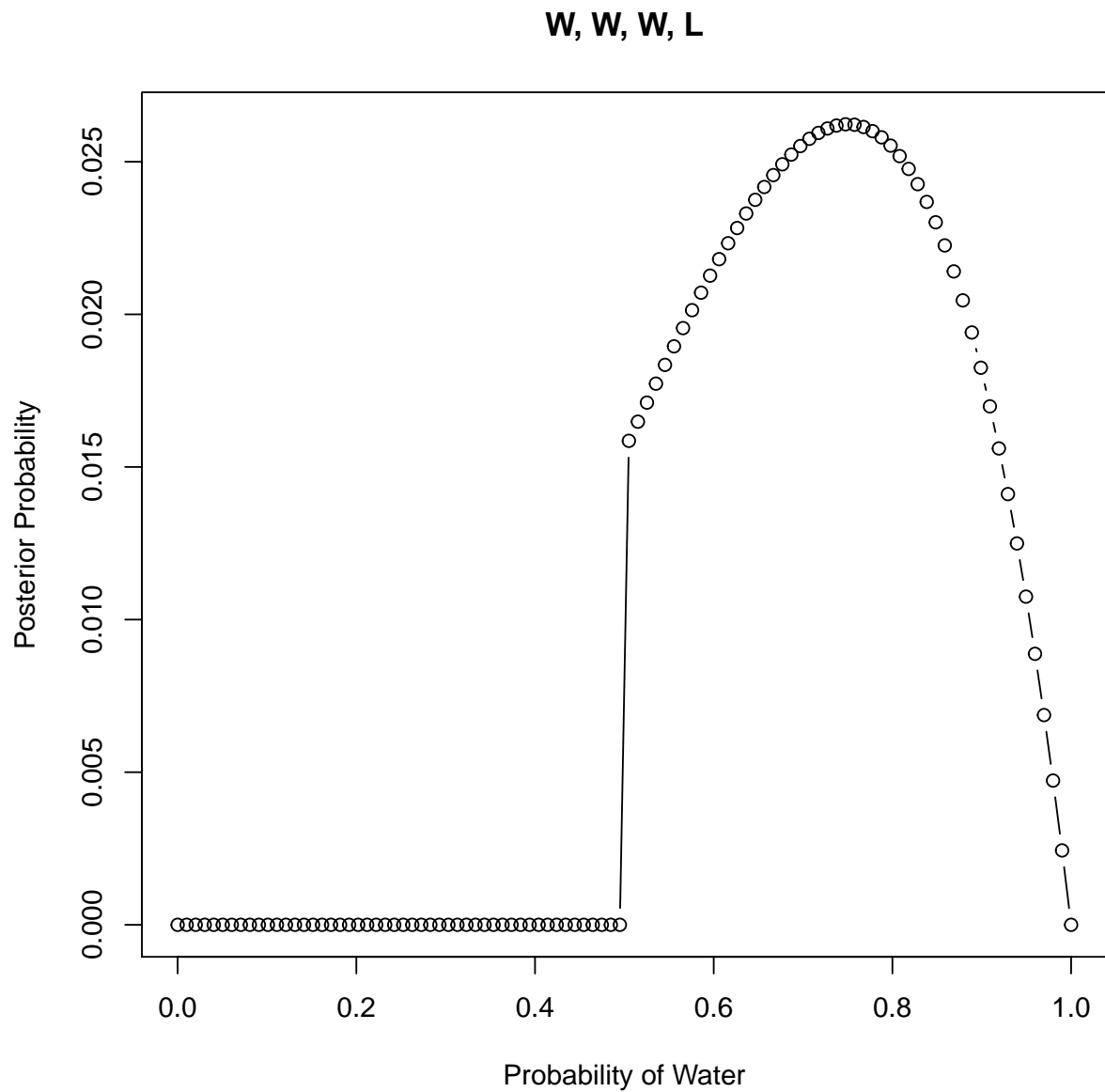
```
#W, W, W, L

likelihood <- dbinom(3, size = 4, prob = p_grid)

unstd.posterior <- likelihood * prior

posterior <- unstd.posterior/sum(unstd.posterior)

plot(p_grid, posterior, type = "b", xlab = "Probability of Water",
      ylab = "Posterior Probability", main = "W, W, W, L")
```



```
#L, W, W, L, W, W, W

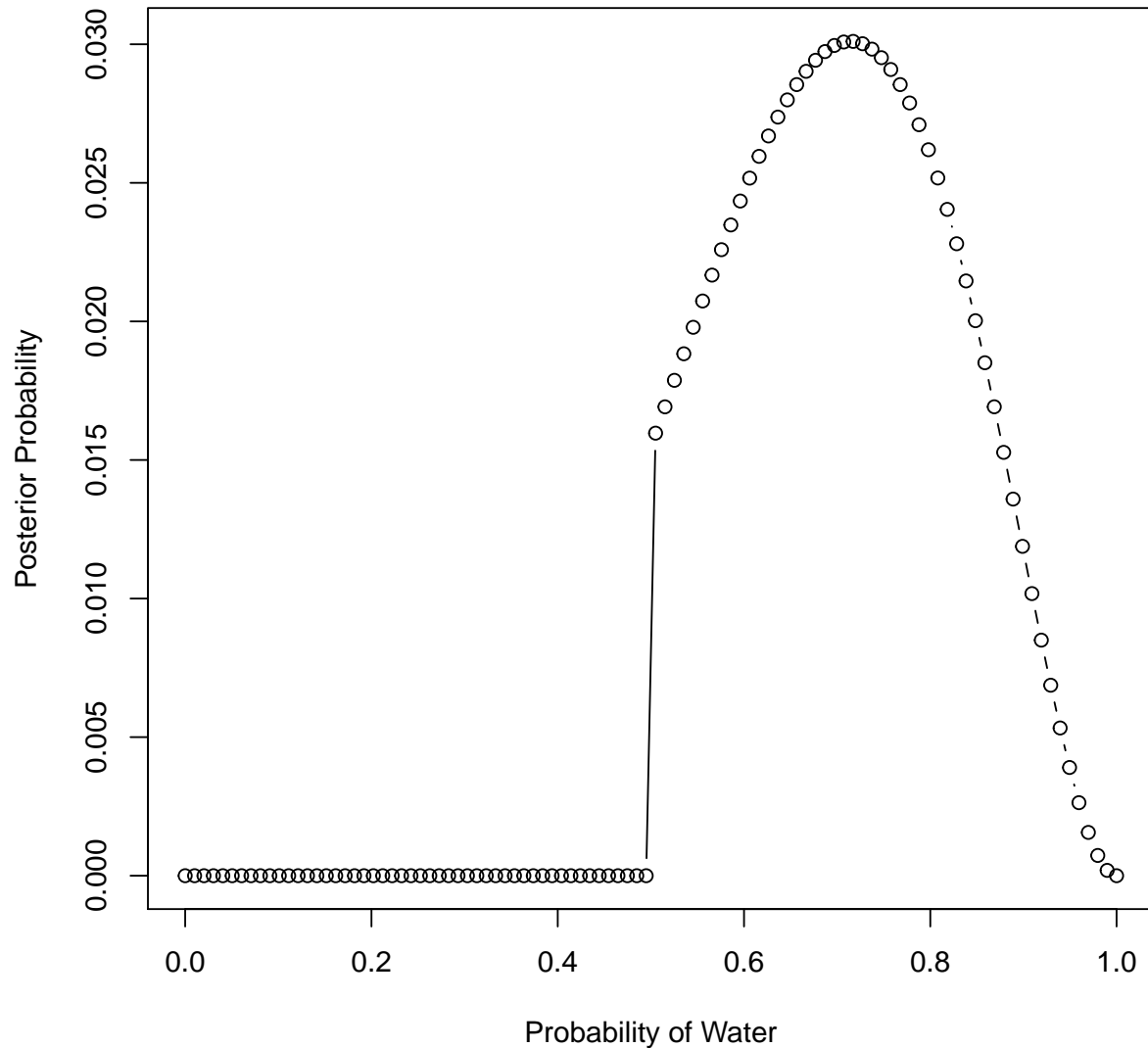
likelihood <- dbinom(5, size = 7, prob = p_grid)

unstd.posterior <- likelihood * prior

posterior <- unstd.posterior/sum(unstd.posterior)

plot(p_grid, posterior, type = "b", xlab = "Probability of Water",
      ylab = "Posterior Probability", main = "L, W, W, L, W, W, W")
```


L, W, W, L, W, W, W



Problem 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" ($\Pr(\text{Earth}|\text{land})$), is 0.23.

$$\Pr(\text{Earth}|\text{land}) = \frac{\Pr(\text{land}|\text{Earth}) \cdot \Pr(\text{Earth})}{\Pr(\text{land})} = \frac{0.3 \cdot 0.5}{0.3 \cdot 0.5 + 1.0 \cdot 0.5} \approx 0.23$$

Problem 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 placed 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 the problem. This means counting up the ways that each card could produce the observed data (a black side facing up on the table).

Scenarios:

1. Up: Card 1, Side 1 (Black), Down: Card 1, Side 2 (Black)
2. Up: Card 1, Side 2 (Black), Down: Card 1, Side 1 (Black)
3. Up: Card 2, Side 1 (Black), Down: Card 2, Side 2 (White)
4. Up: Card 2, Side 2 (White), Down: Card 2, Side 1 (Black)
5. Up: Card 3, Side 1 (White), Down: Card 3, Side 2 (White)
6. Up: Card 3, Side 2 (White), Down: Card 3, Side 1 (White)

There are 3 possible scenarios where the side facing up is black (Scenarios 1, 2, & 3). In two of these three scenarios, the other side of the card is also black. Thus, the probability that the other side is also black is $2/3$.

Problem 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.

Scenarios:

1. Up: Card 1, Side 1 (Black), Down: Card 1, Side 2 (Black)
2. Up: Card 1, Side 2 (Black), Down: Card 1, Side 1 (Black)
3. Up: Card 2, Side 1 (Black), Down: Card 2, Side 2 (White)
4. Up: Card 2, Side 2 (White), Down: Card 2, Side 1 (Black)
5. Up: Card 3, Side 1 (White), Down: Card 3, Side 2 (White)
6. Up: Card 3, Side 2 (White), Down: Card 3, Side 1 (White)
7. Up: Card 4, Side 1 (Black), Down: Card 4, Side 2 (Black)
8. Up: Card 4, Side 2 (Black), Down: Card 4, Side 1 (Black)

There are 5 possible scenarios where the side facing up is black (Scenarios 1, 2, 3, 7, & 8). In four of these five scenarios, the other side of the card is also black. Thus, the probability that the other side is also black is $4/5$.

Problem 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.

Card 1 has 2 ways to produce the data, while Card 2 has 1 way (as shown in the scenarios in the previous questions). Adding in the multipliers, we now have Card 1 with $2*1 = 2$ ways to produce the data and Card 2 with $1*2 = 2$ ways to produce the data. Since the other side of Card 1 is always black and the other side of Card 2 is white if the black side is up, we have:

$$\frac{2*1+0*2}{2*1+1*2} = \frac{2}{4} = \frac{1}{2}$$

Problem 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.

Scenarios:

1. Scenario 1

- Up: Card 1, Side 1 (Black), Down: Card 1, Side 2 (Black)
- Up: Card 2, Side 2 (White), Down: Card 2, Side 1 (Black)

2. Scenario 2

- Up: Card 1, Side 1 (Black), Down: Card 1, Side 2 (Black)
- Up: Card 3, Side 1 (White), Down: Card 3, Side 2 (White)

3. Scenario 3

- Up: Card 1, Side 1 (Black), Down: Card 1, Side 2 (Black)
- Up: Card 3, Side 2 (White), Down: Card 3, Side 1 (White)

4. Scenario 4

- Up: Card 1, Side 2 (Black), Down: Card 1, Side 1 (Black)
- Up: Card 2, Side 2 (White), Down: Card 2, Side 1 (Black)

5. Scenario 5

- Up: Card 1, Side 2 (Black), Down: Card 1, Side 1 (Black)
- Up: Card 3, Side 1 (White), Down: Card 3, Side 2 (White)

6. Scenario 6

- Up: Card 1, Side 2 (Black), Down: Card 1, Side 1 (Black)
- Up: Card 3, Side 2 (White), Down: Card 3, Side 1 (White)

7. Scenario 7

- Up: Card 2, Side 1 (Black), Down: Card 2, Side 2 (White)
- Up: Card 3, Side 1 (White), Down: Card 3, Side 2 (White)

8. Scenario 8

- Up: Card 2, Side 1 (Black), Down: Card 2, Side 2 (White)
- Up: Card 3, Side 2 (White), Down: Card 3, Side 1 (White)

As can be seen, there are 8 separate scenarios where the first card shows a black side and the second card shows a white side. In 6 out of 8 scenarios, the other side of the first card is also black. Thus, the probability that the other side of the first card is also black when the first card is showing black and the second card is showing white is: $\frac{6}{8} = 0.75$.

Problem 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 = 1/2 # prior probability of species A
p_B = 1/2 # prior probability of species B

p_twins_A = 0.1 # P(Twins|Species = A)
p_twins_B = 0.2 # P(Twins|Species = B)
```

```

# Given first birth are twins, what is the probability that
# second birth are also twins

# Given a species, births should be independent.
# But, because we are unsure about the species, we need to
# average out the species to estimate the probability of
# next birth being twins, given that the first birth
# were twins.

# In other words, if I knew the species, knowing that the current
# birth resulted in twins, would not change my expectation about
# twins for the following birth.

# Thus, in principle:
#  $P(\text{twins}) = P(\text{twins}|\text{Species} = A)P(A) + P(\text{twins}|\text{Species} = B)P(B)$ 

# However, knowing about a single twin birth, does change
# the relative plausibilities of the two species. So,  $P(A)$  and
#  $P(B)$ , are actually  $P(A|\text{twins})$  and  $P(B|\text{twins})$ .

# Thus, we are actually looking for:
#  $P(\text{twins}) = P(\text{twins}|A)P(A|\text{twins}) + P(\text{twins}|B)P(B|\text{twins})$ 

#  $P(A|\text{twins}) = P(\text{twins}|A)P(A)/P(\text{twins})$ 
#  $P(B|\text{twins}) = P(\text{twins}|B)P(B)/P(\text{twins})$ 
#  $P(\text{twins}) = P(\text{twins}|A)P(A) + P(\text{twins}|B)P(B)$ 
p_twins_unadjusted = 0.1 * 0.5 + 0.2 * 0.5

p_A_twins = (0.1 * 0.5)/p_twins_unadjusted
p_B_twins = (0.2 * 0.5)/p_twins_unadjusted

p_twins_adjusted = 0.1 * p_A_twins + 0.2 * p_B_twins
print(p_twins_adjusted)

## [1] 0.1666667

```

Thus, $\Pr(\text{twins}) \approx 0.167$.

Problem 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 is twins.

```

#  $P(A|twins) = P(twins|A)P(A)/P(twins)$ 

p_twins = 0.1 * 0.5 + 0.2 * 0.5

p_A_twins = (0.1 * 0.5)/p_twins

print(p_A_twins)

## [1] 0.3333333

```

Thus, $\Pr(A|twins) \approx 0.33$.

Problem 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.

```

# First birth = Twins
# Second birth = Singleton
#  $P(\text{Species} = A | \text{Twins}, \text{Singleton})?$ 

#  $P(\text{Species} = A | \text{Twins}, \text{Singleton}) = P(\text{Twins} | \text{Species} = A) *$ 
#  $P(\text{Singleton} | \text{Species} = A) *$ 
#  $P(A) /$ 
#  $P(\text{Twins}, \text{Singleton})$ 
#  $P(\text{Twins}, \text{Singleton}) = P(\text{Twins} | \text{Species} = A) *$ 
#  $P(\text{Singleton} | \text{Species} = A) *$ 
#  $P(A) +$ 
#  $P(\text{Twins} | \text{Species} = B) *$ 
#  $P(\text{Singleton} | \text{Species} = B) *$ 
#  $P(B)$ 

p_A_twins_singleton = ( 0.1 * 0.9 * 0.5 ) /
    (( 0.1 * 0.9 * 0.5 ) +
     ( 0.2 * 0.8 * 0.5 ))

print(p_A_twins_singleton)

## [1] 0.36

```

Thus, $\Pr(A|twins, singleton) = 0.36$.

Problem 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 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.

```
# Probability of species A given a postive genetic test for A?

# P(test = A | Species = A) = 0.80
# P(test = B | Species = A) = 0.20
# P(test = B | Species = B) = 0.65
# P(test = A | Species = B) = 0.35

# P(Species = A | Test = A) =
#   P(Test = A | Species = A) * P(Species = A) / P(Test = A)

#P(test = A) =
#   P(test = A | Species = A) * P(Species = A) +
#   P(test = A | Species = B) * P(Species = B)

p_a_testA = 0.8 * 0.5 / (0.8*0.5+0.35*0.5)

print(p_a_testA)

## [1] 0.6956522
```

Thus, $\Pr(A|\text{Test} = A) \approx 0.696$.

```
# P(Species = A | test = A, twins, singleton)?

# P(Species = A | test = A, twins, singleton) =
#   P(test = A | Species = A) *
#   P(twins | Species = A) *
#   P(singleton | Species = A) *
#   P(Species = A) \
#   P(test = A, twins, singleton)

# P(test = A, twins, singleton) =
#   P(test = A | Species = A) *
```

```

#           P(twins | Species = A) *
#           P(singleton | Species = A) *
#           P(Species = A) +
#           P(test = A | Species = B) *
#           P(twins | Species = B) *
#           P(singleton | Species = B) *
#           P(Species = B)

# The individual probabilities are only conditional on
# species being equal to A because otherwise they are
# independent. A test = A does not depend on births, and
# births should be independent.

p_testA_twins_singleton = 0.8 * 0.1 * 0.9 * 0.5 +
                          0.35 * 0.2 * 0.8 * 0.5

p_A_testA_twins_singleton = (0.8 * 0.1 * 0.9 * 0.5) / p_testA_twins_singleton

print(p_A_testA_twins_singleton)

## [1] 0.5625

```

Thus, $\Pr(A | \text{Test} = A, \text{twins}, \text{singleton}) = 0.5625$.