Extended Practice

Chapter 2, Lab 4: Solutions

OpenIntro Biostatistics

This lab contains examples of relatively complex conditioning problems that require particular care in both translating problem statements to the language of probability and correctly manipulating marginal, joint, and conditional probabilities.

The first problem is from the context of allele inheritance, like Examples 2.4 and 2.32 in the text. The second problem is approached algebraically in the text in Section 2.3; a familiarity with both algebraic and simulation methods is instructive.

Problem 1: Eye Color Inheritance.

One of the earliest models for the genetics of eye color was developed in 1907, and proposed a single-gene inheritance model, for which brown eye color is always dominant over blue eye color.[1]

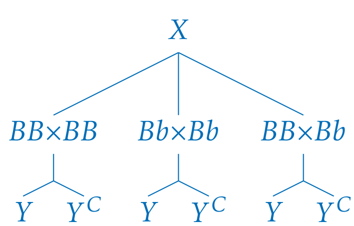
Suppose that in the population, 25% of individuals are homozygous dominant (genotype *BB*), 50% are heterozygous (genotype *Bb*), and 25% are homozygous recessive (genotype *bb*). Individuals with genotype *BB* or *Bb* have brown eyes, while those with genotype *bb* have blue eyes.

Assume independent mating in the population.

a) Suppose that two parents have brown eyes. What is the probability that their first child has blue eyes?

   Let X represent the event of two parents having brown eyes, and Y represent the event of having a child with blue eyes. The following tree illustrates the possible paths to Y.

   Note that the $\times$ represents a mating between two genotypes; e.g. $BB \times BB$ represents a mating between two BB individuals. The symbol $\cdot$ is used for multiplication of probabilities



   Solve for $P(Y|X)$, which is given by adding together the probabilities of the three branches ending in Y.[2]

   Find the probabilities of the first set of branches: $P(BB \times BB|X)$, $P(Bb \times Bb|X)$, and $P(BB \times Bb|X)$. In other words, find the probability of each pair, given that both parents have brown eyes.

[1] This model has since been shown to be incorrect—eye color inheritance is actually controlled by multiple genes.

[2] The total probability equation : $P(Y|X) = [P(BB \times BB|X) \cdot P(Y|BB \times BB)] + [P(Bb \times Bb|X) \cdot P(Y|Bb \times Bb)] + [P(BB \times Bb|X) \cdot P(Y|BB \times Bb)]$

- Calculate $P(BB \times BB|X)$. Assuming independent mating, the probability of two individuals of genotype BB mating, given that they both have brown eyes, is equal to $[P(BB|X)]^2$.

  To find $P(BB|X)$, use Bayes' rule with the genotype frequencies given in the population. An individual can have brown eyes from either being BB or Bb.

$$\begin{aligned} P(BB \times BB|X) = [P(BB|X)]^2 &= \left[\frac{P(BB \text{ and } X)}{P(X)}\right]^2 \\ &= \left[\frac{P(X|BB) \cdot P(BB)}{[P(X|BB) \cdot P(BB)] + [P(X|Bb) \cdot P(Bb)]}\right]^2 \\ &= \left[\frac{(1)(1/4)}{[(1)(1/4)] + [(1)(1/2)]}\right]^2 \\ &= \left[\frac{1}{3}\right]^2 \\ &= \frac{1}{9} \end{aligned}$$

- Calculate $P(Bb \times Bb|X)$.

$$\begin{aligned} P(Bb \times Bb|X) = [P(Bb|X)]^2 &= \left[\frac{P(Bb \text{ and } X)}{P(X)}\right]^2 \\ &= \left[\frac{P(X|Bb) \cdot P(Bb)}{[P(X|BB) \cdot P(BB)] + [P(X|Bb) \cdot P(Bb)]}\right]^2 \\ &= \left[\frac{(1)(1/2)}{[(1)(1/4)] + [(1)(1/2)]}\right]^2 \\ &= \left[\frac{2}{3}\right]^2 \\ &= \frac{4}{9} \end{aligned}$$

Note that it is also possible to find $P(BB|X)$ and $P(Bb|X)$ from applying the definition of conditional probability and using the genotype frequencies:

$$P(BB \times BB|X) = [P(BB|X)]^2 = \left[\frac{P(BB \text{ and } X)}{P(X)}\right]^2 = \left[\frac{0.25}{0.75}\right] = (1/3)^2 = 1/9$$

$$P(Bb \times Bb|X) = [P(Bb|X)]^2 = \left[\frac{P(Bb \text{ and } X)}{P(X)}\right]^2 = \left[\frac{0.50}{0.75}\right] = (2/3)^2 = 4/9$$

- Calculate $P(BB \times Bb|X)$. From the previous calculations, $P(BB|X)$ and $P(Bb|X)$ are already known to be $1/3$ and $2/3$, respectively. Since the mother could be BB and the father Bb, or vice versa, remember to multiply by 2.

$$
\begin{aligned}
P(BB \times Bb|X) =& 2 \cdot P(BB|X) \cdot P(Bb|X) \\
=& (2)(1/3)(2/3) \\
=& \frac{4}{9}
\end{aligned}
$$

Find the probabilities of the second set of branches: $P(Y|BB \times BB)$, $P(Y|Bb \times Bb)$, and $P(Y|BB \times Bb)$. In other words, find the probability of a child with blue eyes, given a specific genotype pair.

A parent passes down one of two alleles (B or b) with equal chance, and alleles are inherited independently of each other. An individual must be genotype bb in order to have blue eyes.

- Calculate $P(Y|BB \times BB)$. These parents can only produce BB children; thus, $P(Y|BB \times BB) = 0$.
- Calculate $P(Y|Bb \times Bb)$. These parents produce bb children with probability $(1/2)(1/2) = 1/4$; $P(Y|Bb \times Bb) = 1/4$.
- Calculate $P(Y|BB \times Bb)$. Since one parent is BB, every child necessarily receives a B allele; $P(Y|BB \times BB) = 0$.

Add together the probabilities of the three branches ending in Y.

$$
\begin{aligned}
P(Y|X) =& [P(BB \times BB|X) \cdot P(Y|BB \times BB)] + [P(Bb \times Bb|X) \cdot P(Y|Bb \times Bb)] + [P(BB \times Bb|X) \cdot P(Y|BB \times Bb)] \\
=& [(1/9)(0)] + [(4/9)(1/4)] + [(4/9)(0)] \\
=& \frac{1}{9}
\end{aligned}
$$

Given that two parents have brown eyes, the probability of their first child having blue eyes is 1/9, or about 0.11.

b) Does the probability in part a) change if it is now known that the paternal grandfather had blue eyes? Justify your answer.

Since the father has brown eyes, with a parent that has blue eyes (genotype bb), he must be genotype Bb, since he necessarily inherited one of his father's b alleles.

This changes $P(Bb \times Bb|X)$ from the previous solution. While $P(Bb|X)$ remains the same for the mother (2/3), $P(Bb|X, \text{pat. grandfather } bb)$ for the father is 1.

$$
\begin{aligned}
P(Bb \times Bb|X, \text{pat. gf } bb) &= P(\text{father } Bb|\text{X, pat. gf } bb) \times P(\text{mother } Bb\ |X, \text{pat. gf } bb) \\
&= 1 \times \frac{2}{3} = \frac{2}{3}
\end{aligned}
$$

Substitute this value into the final equation from part a):

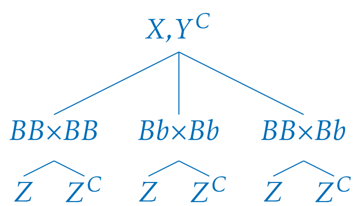
$$\begin{aligned}P(Y|X) =& 0 + [P(Bb \times Bb|X) \cdot P(Y|Bb \times Bb)] + 0\\ =&(2/3)(1/4)\\ =&\frac{1}{6}\end{aligned}$$

The probability has changed from 1/9 to 1/6. Note that the probability is larger, which makes sense since the given information increases the probability that the father is Bb, rather than BB (and BB individuals cannot have blue-eyed children).

c) Given that their first child has brown eyes, what is the probability that their second child has blue eyes? Ignore the condition given in part b).

This question builds on the scenario presented in part a), conditioning on the event that the first child has brown eyes in addition to the event that both parents have brown eyes.

Let Y^C represent the event of having a first child with brown eyes, and Z represent the event of having a second child with blue eyes.



Solve for $P(Z|X, Y^C)$.

Start by writing out the total probability equation.

$$\begin{aligned}P(Z|X, Y^C) = &[P(BB \times BB|X, Y^C) \cdot (P(Z|BB \times BB, X, Y^C)]\\ &+ [P(Bb \times Bb|X, Y^C) \cdot (P(Z|Bb \times Bb, X, Y^C)]\\ &+ [P(BB \times Bb|X, Y^C) \cdot (P(Z|BB \times Bb, X, Y^C)]\end{aligned}$$

The second part of each term, the probability of Z given a combination of parental genotypes (and events X and Y^C), can be calculated from the rules of genetics. Only a $Bb \times Bb$ pairing can potentially produce a blue-eyed child; $(P(Z|BB{\times}BB, X, Y^C) = 0$, $(P(Z|Bb{\times}Bb, X, Y^C) = 1/4$, $(P(Z|BB \times Bb, X, Y^C) = 0$.

Substitute these values into the total probability equation and simplify.

$$\begin{aligned}P(Z|X, Y^C) &= [P(BB \times BB|X, Y^C)(0)] + [P(Bb \times Bb|X, Y^C)(1/4)] + [P(BB \times Bb|X, Y^C)(0)]\\ &= P(Bb \times Bb|X, Y^C)(1/4)\end{aligned}$$

Find $P(Bb \times Bb|X, Y^C)$. This probability can be found using Bayes' Rule, with Y^C being partitioned into the three different ways it can occur (i.e. the three possible genotype pairs).

$$P(Bb \times Bb|X, Y^C) = \frac{P(Bb \times Bb|X) \cdot P(Y^C|Bb \times Bb, X)}{P(Y^C|X)}$$

Note that the denominator, $P(Y^C|X)$, is the complement of the probability solved for in part a), $P(Y|X)$. Thus, $P(Y^C|X) = 1 - 1/9 = 8/9$. Alternatively, fully expand the denominator to solve for $P(Y^C|X)$:

$$\begin{aligned}
P(Y^C|X) &= [(P(BB \times BB|X)) \cdot (P(Y^C|BB \times BB, X))] + [(P(BB \times Bb|X)) \cdot (P(Y^C|BB \times Bb, X))] \\
&\quad + [(P(Bb \times Bb|X)) \cdot (P(Y^C|Bb \times Bb, X))] \\
&= \left[\left(\frac{1}{3}\right)^2 \times 1\right] + \left[(2)\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)(1)\right] + \left[\left(\frac{2}{3}\right)^2 \frac{3}{4}\right] \\
&= \frac{1}{9} + \frac{4}{9} + \frac{3}{9} \\
&= \frac{8}{9}
\end{aligned}$$

Finish solving for $P(Bb \times Bb|X, Y^C)$:

$$\begin{aligned}
P(Bb \times Bb|X, Y^C) &= \frac{P(Bb \times Bb|X) \cdot P(Y^C|Bb \times Bb, X)}{P(Y^C|X)} \\
&= \frac{\left[\left(\frac{2}{3}\right)^2 \left(\frac{3}{4}\right)\right]}{\frac{8}{9}} = \frac{3}{8}
\end{aligned}$$

Return to the simplified total probability equation.

$$\begin{aligned}
P(Z|X, Y^C) &= P(Bb \times Bb|X, Y^C)(1/4) \\
&= (3/8)(1/4) \\
&= 3/32 = 0.09375
\end{aligned}$$

Problem 2: Cat Genetics.

The gene that controls white coat color in cats, KIT, is known to be responsible for multiple phenotypes such as deafness and blue eye color. A dominant allele W at one location in the gene has complete penetrance for white coat color; all cats with the W allele have white coats. There is incomplete penetrance for blue eyes and deafness; not all white cats will have blue eyes and not all white cats will be deaf. However, deafness and blue eye color are strongly linked, such that white cats with blue eyes are much more likely to be deaf. The variation in penetrance for eye color and deafness may be due to other genes as well as environmental factors.

Suppose that 30% of white cats have one blue eye, while 10% of white cats have two blue eyes. About 73% of white cats with two blue eyes are deaf and 40% of white cats with one blue eye are deaf. Only 19% of white cats with other eye colors are deaf.

Create a simulated dataset of 100,000 white cats to answer the following questions. The lab template provides a basic outline of the components of the simulation.

The structure of the simulation is essentially identical to that used in the two previous labs. Eye color is assigned using only the `sample()` command, while deafness and blindness are assigned with `if()` statements. Note how blindness must be assigned conditional on both deafness and eye color.

```r
#define parameters
population.size = 100000

blue.one = 0.30
blue.two = 0.10
blue.zero = 1 - blue.one - blue.two

deaf.of.blue.two = 0.73
deaf.of.blue.one = 0.40
deaf.of.blue.zero = 0.19

blind.of.deaf.blue.two.zero = 0.20
blind.of.deaf.blue.one = 0.40
blind.of.not.deaf = 0.10

#create empty vectors to store results
blue.eyes = vector("numeric", population.size)
deafness = vector("numeric", population.size)
blindness = vector("numeric", population.size)

#set the seed for a pseudo-random sample
set.seed(2018)

#assign eye color
blue.eyes = sample(c(0,1,2), size = population.size,
                   prob = c(blue.zero, blue.one, blue.two),
                   replace = TRUE)
```

```r
#assign deafness
for(k in 1:population.size){

  if(blue.eyes[k] == 0){
    deafness[k] = sample(c(0,1), size = 1,
                          prob = c(1 - deaf.of.blue.zero, deaf.of.blue.zero))
  }

  if(blue.eyes[k] == 1){
    deafness[k] = sample(c(0,1), size = 1,
                          prob = c(1 - deaf.of.blue.one, deaf.of.blue.one))
  }

  if(blue.eyes[k] == 2){
    deafness[k] = sample(c(0,1), size = 1,
                       prob = c(1 - deaf.of.blue.two, deaf.of.blue.two))
  }

}

#assign blindness

for(k in 1:population.size){

  if((blue.eyes[k] == 0 & deafness[k] == 1) | (blue.eyes[k] == 2 & deafness[k] == 1)){
    blindness[k] = sample(c(0,1), size = 1,
                        prob = c(1 - blind.of.deaf.blue.two.zero,
                                 blind.of.deaf.blue.two.zero))
  }

  if (blue.eyes[k] == 1 & deafness[k] == 1){
    blindness[k] = sample(c(0,1), size = 1,
                        prob = c(1 - blind.of.deaf.blue.one,
                                 blind.of.deaf.blue.one))


  }

  if (deafness[k] == 0){
    blindness[k] = sample(c(0,1), size = 1,
                        prob = c(1 - blind.of.not.deaf,
                                 blind.of.not.deaf))
  }

}
```

a) Calculate the prevalence of deafness among white cats.

The prevalence of deafness among white cats is 0.307.

```
sum(deafness)/population.size
```

```
## [1] 0.30659
```

b) Given that a white cat is deaf, what is the probability that it has two blue eyes?

The probability that a white cat has two blue eyes given that it is deaf is 0.237.

```
sum(blue.eyes == 2 & deafness == 1)/sum(deafness == 1)
```

```
## [1] 0.2374507
```

c) Suppose that deaf, white cats have an increased chance of being blind, but that the prevalence of blindness differs according to eye color. While deaf, white cats with two blue eyes or two non-blue eyes have probability 0.20 of developing blindness, deaf and white cats with one blue eye have probability 0.40 of developing blindness. White cats that are not deaf have probability 0.10 of developing blindness, regardless of their eye color.

i. What is the prevalence of blindness among deaf, white cats?

The prevalence of blindness among deaf, white cats is 0.276.

```
sum(blindness == 1 & deafness == 1)/sum(deafness == 1)
```

```
## [1] 0.2764278
```

ii. What is the prevalence of blindness among white cats?

The prevalence of blindness among white cats is 0.155.

```
sum(blindness)/population.size
```

```
## [1] 0.15529
```

iii. Given that a cat is white and blind, what is the probability that it has two blue eyes?

The probability a cat has two blue eyes given it is white and blind is 0.113.

```
sum(blue.eyes == 2 & blindness == 1)/sum(blindness == 1)
```

```
## [1] 0.1128212
```

iv. What is the probability that a white cat has one blue eye and one non-blue eye, given that it is not blind?

The probability that a cat with one blue eye and one non-blue eye, given that it is not blind, is 0.276. The algebraic solution to this problem is shown as Example 2.58 in the text.

```
sum(blue.eyes == 1 & blindness == 0)/sum(blindness == 0)
```

```
## [1] 0.2755975
```

d) The following two questions are not solved in the text. Solve the answers algebraically, then check the answers with the results of the simulation.

i. What is the prevalence of two blue eyes among white cats that are both deaf and blind?

Let D represent the event that a white cat is deaf. Let B_2 represent the event that a white cat has two blue eyes. Let L represent the event that a white cat is blind.

$$\begin{aligned}P(B_2|L,D) =& \frac{P(L \text{ and } D \text{ and } B_2)}{P(L \text{ and } D)}\\ =& \frac{P(L|D,B_2)P(D \text{ and } B_2)}{P(L \text{ and } D)}\\ =& \frac{P(L|D,B_2)P(B_2|D)P(D)}{P(L|D)P(D)}\\ =& \frac{(0.20)(0.238)(0.307)}{(0.278)(0.307)}\\ =& 0.171\end{aligned}$$

```
sum(blue.eyes == 2 & blindness == 1 & deafness == 1)/sum(blindness == 1 & deafness == 1)
```

```
## [1] 0.1729794
```

ii. Given that a white cat has two blue eyes, what is the probability that it is deaf and blind?

$$\begin{aligned}P(L \text{ and } D|B_2) =& \frac{P(L \text{ and } D \text{ and } B_2)}{P(B_2)}\\ =& \frac{P(L|D,B_2)P(D \text{ and } B_2)}{P(B_2)}\\ =& \frac{P(L|D,B_2)P(B_2|D)P(D)}{P(B_2)}\\ =& \frac{(0.20)(0.238)(0.307)}{0.10}\\ =& 0.146\end{aligned}$$

```
sum(blue.eyes == 2 & blindness == 1 & deafness == 1)/sum(blue.eyes == 2)
```

```
## [1] 0.1465267
```