Homework 2

Autumn 2022

Jaiden Atterbury

Instructions

- This homework is due in Gradescope on Wednesday Oct 19 by midnight PST.
- Please answer the following questions in the order in which they are posed.
- Don't forget to knit the document frequently to make sure there are no compilation errors.
- When you are done, download the PDF file as instructed in section and submit it in Gradescope.

Exercises

1. Suppose 12 coins are tossed and the outcome (head or tail) is recorded for each.

a. The sample space S corresponding to this "experiment" consists of all possible sequences of heads and tails that result from tossing 12 coins. How many elements are in S? Calculate this number in a code chunk and also explain your answer very briefly.

   Hint: Refer to Braille alphabet example 3.2. This problem is similar but with a 12 dot matrix.

$S = \{$all possible sequences of heads and tails that result from tossing 12 coins.$\}$

```
num_in_S <- 2^12
```

Using the multiplication principle for counting, this "job" consists of 12 seperate tasks peformed in series. The i^{th} of which can be done in 2 seperate ways. Therefore, there are $|S| = 2^{12} = 4096$ elements in S.

b. Let E denote the event that 7 of the 12 coins land on heads. How many elements are in E? Explain your answer very briefly.

   Hint: Referring to the Braille alphabet example again, think about counting the number of letters that can be formed with 3 raised dots. What decision do you need to make? How many ways can you make it?

$E = \{$7 of the 12 coins land on heads.$\}$

```
num_in_E <- choose(12,7)
```

Since we are only interested in the number of groups of 7 items that can be formed from 12, position doesn't matter. Thus, we can use a binomial coefficent to find the humber of elements in E. In this case $|E| = \binom{12}{7} =$ 792 elements in E.

c. Calculate $P(E)$ assuming all the elements in S are equally likely. Report your final answer in a sentence using inline code.

```
prob_of_E <- num_in_E / num_in_S
```

Using the equally likely rule, we find that the $P(E) = \frac{|E|}{|S|} = \frac{792}{4096} = 0.1933594$.

2. To estimate the number N of goldfish in a pond, $r = 25$ fish were caught, tagged and released. Later, a second sample of $n = 20$ fish were caught and 5 fish in this sample were noted to be tagged.

a. How many possible samples of size $n = 20$ can be formed from the N fish in the pond? (Leave your answer in terms of a binomial coefficient - you cannot calculate it because you don't know N)

There are $|S| = \binom{N}{20}$ possible samples of size $n = 20$ that can be formed from the N fish in the pond.

b. The event E contains all the samples which have 5 tagged and 15 untagged fish. How many elements are in the event E? (Leave your answer in terms of N)

$E = \{$all the samples which have 5 tagged and 15 untagged fish.$\}$

The event E consists of all samples of 5 tagged fish that can be chosen from the 25 that were intitially tagged in the first sample, and all of the samples of 15 untagged fish that can be chosen from the $N - 25$ untagged fish remaining in the pond. Thus, $|E| = \binom{25}{5} * \binom{N-25}{15}$.

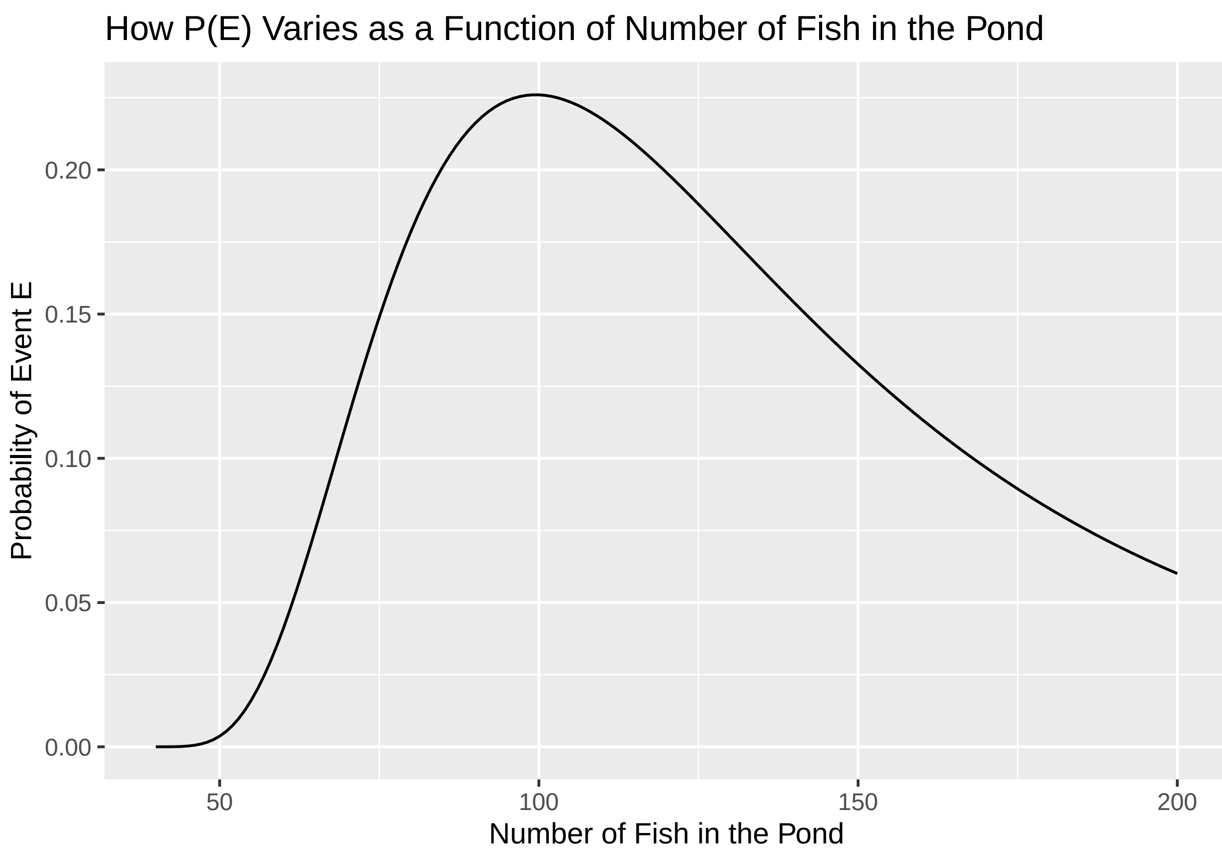
c. Assuming each possible sample is equally likely, give an expression for $P(E)$. (Leave your answer in terms of N)

Using the equally likely rule, we find that $P(E) = \frac{|E|}{|S|} = \frac{\binom{25}{5} * \binom{N-25}{15}}{\binom{N}{20}}$.

d. In this part, we will examine visually how $P(E)$ varies as a function of N. Fill in the blanks in the R code provided and run it in R to create the plot. You should remove the `eval = FALSE` chunk option before knitting. (Note: this part does not require you to know **tidyverse** or **ggplot** functions, and should be doable by all.)

```
fishes <- tibble( #enhanced data frame
        N = 40:200,     #possible values for N: 40,41, ...,200
        prob = (choose(25, 5) * choose(N - 25, 15)) / (choose(N, 20))
               )  #write expression for P(E) in terms of N

ggplot(data = fishes,
       mapping = aes(x = N, y = prob)) +
           geom_line() +
             labs(title = "How P(E) Varies as a Function of Number of Fish in the Pond",
                  x = "Number of Fish in the Pond",
                  y = "Probability of Event E")
```

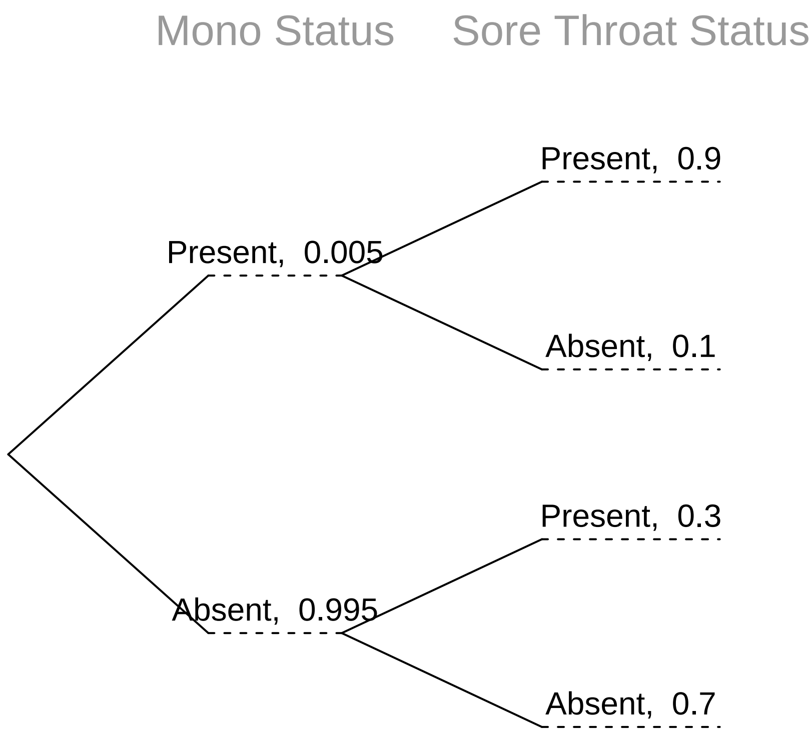


3. Among all students seeking treatment at Hall Health, 0.5% are eventually diagnosed as having mononucleosis (event A). Of those who do have mono, 90% complain of a sore throat (event B). But 30% of those not having mono also have sore throats.

a. Make a tree diagram of the probabilities relating presenting with a sore throat and a diagnosis of mononucleosis. (Don't forget to include the **openintro** package in the setup chunk)

Calculating the probabilities for the tree diagram:
$P(A) = 0.005$
$P(B|A) = 0.9$
$P(B^c|A) = 1 - P(B|A) = 1 - 0.9 = 0.1$

$P(A^c) = 1 - P(A) = 1 - 0.005 = 0.995$
$P(B|A^c) = 0.3$
$P(B^c|A^c) = 1 - (B|A^c) = 1 - 0.3 = 0.7$



b. If a student comes to Hall Health and says that she has a sore throat, what is the probability that she has mono? Be sure to show your steps carefully.

```
P_A_and_B <- 0.005 * 0.9
P_Ac_and_B <- 0.995 * 0.3
P_B <- P_A_and_B + P_Ac_and_B
P_A_given_B <- P_A_and_B / P_B
```

This question is asking us to find $P(A|B) = \frac{P(A \cap B)}{P(B)}$. In order to find the denominator: the event, B, that the the student has a sore throat is the union of two mutually exclusive events: $B = (B \cap A) \cup (B \cap A^c)$. Since $P(B \cap A) = 0.0045$, and $P(B \cap A^c) = 0.2985$, it follows that, $P(B) = 0.303$. Therefore, $P(A|B) = \frac{0.0045}{0.303} = 0.0148515$.

4. Prove that if A and B are independent events, then the following pairs are also independent.

Hint: In each part, you are trying to show the two events in question satisfy the definition of independence. To prove this for a, write the set A as the union of two disjoint events $A \cap B$ and $A \cap B^c$, then use axiom 3, followed by knowledge of independence of A and B and see if that gets you anything.

a. A and B^c

Proof:

$$\begin{aligned}
A &= (A \cap B) \cup (A \cap B^c) \\
P(A) &= P(A \cap B) + P(A \cap B^c) \quad \text{(Axiom 3)} \\
P(A) &= P(A) * P(B) + P(A \cap B^c) \quad \text{(Definition 4.1)} \\
P(A \cap B^c) &= P(A) - P(A) * P(B) \quad \text{(Isolating } P(A \cap B^c)\text{)} \\
P(A \cap B^c) &= P(A) * (1 - P(B)) \quad \text{(Factoring P(A))} \\
P(A \cap B^c) &= P(A) * P(B^c) \quad \text{(Theorem 2.1.a.)}
\end{aligned}$$

Therefore, by Definition 4.1, A and B^c are independent.

b. A^c and B

Proof:

$$\begin{aligned}
B &= (B \cap A) \cup (B \cap A^c) \\
P(B) &= P(B \cap A) + P(B \cap A^c) \quad \text{(Axiom 3)} \\
P(B) &= P(B) * P(A) + P(B \cap A^c) \quad \text{(Definition 4.1)} \\
P(B \cap A^c) &= P(B) - P(B) * P(A) \quad \text{(Isolating } P(B \cap A^c)\text{)} \\
P(B \cap A^c) &= P(B) * (1 - P(A)) \quad \text{(Factoring P(B))} \\
P(B \cap A^c) &= P(B) * P(A^c) \quad \text{(Theorem 2.1.a.)}
\end{aligned}$$

Therefore, by Definition 4.1, B and A^c are independent.

c. A^c and B^c.

Proof:

$$\begin{aligned}
P(A^c \cap B^c) &= P((A \cup B)^c) \quad \text{(DeMorgan's Law)} \\
&= 1 - P(A \cup B) \quad \text{(Theorem 2.1.a.)} \\
&= 1 - P(A) - P(B) + P(A \cap B) \quad \text{(Theorem 2.2)} \\
&= 1 - P(A) - P(B) + P(A) * P(B) \quad \text{(Definition 4.1)} \\
&= (1 - P(A)) * (1 - P(B)) \\
&= P(A^c) * P(B^c) \quad \text{(Theorem 2.1.a.)}
\end{aligned}$$

Therefore, by Definition 4.1, A^c and B^c are independent.