section 2.2 Bayes’ Theorem

1. Reconsider the legal case that we analyzed in class. Recall that a forensic expert testified that only 0.32% of the general population had the blood characteristics of the assailant, and of course the defendant had those blood characteristics. We used Bayes’ Theorem to express the conditional probability of guilt, in light of this forensic evidence, as a function of the prior probability of guilt, prior to seeing this forensic evidence. When we evaluated this conditional probability of guilt with a prior probability of .5, we obtained .9968 for this conditional probability of guilt.

Investigate whether it’s a general result or just a coincidence (or some other explanation) that the updated probability of guilt, starting from a prior probability of .5, appears to equal one minus the probability reported for the general population to have the blood characteristics.

More specifically: Let $L$ be the proportion of the general population that has the blood characteristics of a criminal, and suppose that the defendant in a trial has those blood characteristics. Is the conditional probability of guilt, from a prior probability of .5, always equal to $1 - L$? If not, can you make any statement about how that conditional probability relates to $L$, and was it just a fluke that this result appears to hold in the particular legal case that we analyzed?

Write a paragraph summarizing your response. Include appropriate calculations and/or derivations to justify your response.

The conditional probability of guilt, from a prior guilt probability of .5, is not in general equal to $1 - L$. In this example it appears to equal $1 - L$ but only does so when rounding to 4 decimal places.

Bayes’ Theorem establishes that \( \Pr(G \mid E) = \frac{\Pr(E \mid G)}{\Pr(E \mid G) \Pr(G) + \Pr(E \mid G^c) \Pr(G^c)} \).

With a prior guilt probability \( \Pr(G) = .5 \), and assuming that the defendant has the blood characteristics, and using $L$ for \( \Pr(E \mid G^c) \), this becomes:

\[
\Pr(G \mid E) = \frac{(1)(.5)}{[(1)(.5) + L(.5)]} = .5 / (.5 + .5L) = 1 / (1 + L).
\]

This result is always true.

The conjecture that \( \Pr(G \mid E) = 1 - L \) is therefore only true when $1 - L = 1 / (1 + L)$, which is when $L = 0$. So, the conjecture is approximately true whenever the blood characteristics are quite rare, with a probability close to 0.
2. Reconsider the class example about estimating the number of fish in a pond using a capture/recapture technique. Recall the data collection process: We capture 10 fish, tag them, release them back into the pond, allow sufficient time for them to mix/mingle thoroughly with the other fish, and then capture 20 fish. Again suppose that we find 4 tagged fish in this second sample. Again let \( N \) represent the total number of fish in the lake.

a) Determine (as we did in class) the probability of obtaining these data (4 tagged fish in the second sample) as a function of \( N \). Graph this function of \( N \), and identify the value(s) of \( N \) that maximize(s) this function.

This probability is \( \frac{\binom{10}{4} \binom{N-10}{16}}{\binom{N}{20}} \) for \( N = 26, 27, \ldots \)

This probability function is maximized at \( N = 49 \) and \( N = 50 \), as shown in the graph below:

![Graph showing probability function]

b) Now suppose that prior to collecting the data, you believe that \( N \) is equally likely to be 20, 30, 40, 50, or 60. Use Bayes’ Theorem to determine the posterior probabilities that \( N = 20, 30, 40, 50, 60 \), conditional on the observed data. Given the data, which of the five possibilities is most likely, and which is least likely?

Let \( E \) represent the observed data (evidence) that 4 of the fish caught in the second sample were tagged.

The denominator for Bayes’ Theorem is:

\[
(1/5)(0) + (1/5)\left(\frac{\binom{10}{4} \binom{30-10}{16}}{\binom{30}{20}}\right) + (1/5)\left(\frac{\binom{10}{4} \binom{40-10}{16}}{\binom{40}{20}}\right) + (1/5)\left(\frac{\binom{10}{4} \binom{50-10}{16}}{\binom{50}{20}}\right) + (1/5)\left(\frac{\binom{10}{4} \binom{60-10}{16}}{\binom{60}{20}}\right) \approx .1564
\]

Pr(\( N = 20 \mid E \)) = 0 because Pr(\( E \mid N = 20 \)) = 0 (least likely)

Pr(\( N = 30 \mid E \)) = \( (1/5)\left(\frac{\binom{10}{4} \binom{30-10}{16}}{\binom{30}{20}}\right) / \text{denom} \approx .043 \)

Pr(\( N = 40 \mid E \)) = \( (1/5)\left(\frac{\binom{10}{4} \binom{40-10}{16}}{\binom{40}{20}}\right) / \text{denom} \approx .283 \)
Pr(N = 50 | E) = (1/5) \left( \frac{10 \choose 4} {50 \choose 16} / \frac{4 \choose 4} {20 \choose 16} \right) / \text{denom} \approx .358 \; (\text{most likely})

Pr(N = 60 | E) = (1/5) \left( \frac{10 \choose 4} {60 \choose 16} / \frac{4 \choose 4} {20 \choose 16} \right) / \text{denom} \approx .315

c) Now suppose that prior to collecting the data, you believe that $N$ is 20, 30, 40, 50, or 60, but you believe that 30 is twice as likely as 20, 40 is twice as likely as 30, 50 is exactly as likely as 30, and 60 is exactly as likely as 20. Determine these prior probabilities. Then use Bayes’ Theorem to determine the posterior probabilities conditional on the observed data. Given the data, which of the five possibilities is most likely, and which is least likely?

Now to make the prior probabilities sum to one, they must be:
Pr(N = 20) = .10, Pr(N = 30) = .20, Pr(N = 40) = .40, Pr(N = 50) = .20, Pr(N = 60) = .10

Applying Bayes’ Theorem in the same way but with these new prior probabilities gives the following posterior probabilities:

Pr(N = 20 | E) = 0 (least likely)
Pr(N = 30 | E) \approx .038
Pr(N = 40 | E) \approx .503 (most likely)
Pr(N = 50 | E) \approx .318
Pr(N = 60 | E) \approx .140

d) Comment on how the posterior probabilities differ for the two different sets of prior probabilities in b) and c).

Because the prior probabilities in (c) put much more weight on $N = 40$, the most likely possibility after seeing the data is now $N = 40$ rather than $N = 50$.

e) Now suppose that prior to collecting the data, you believe that $N$ is equally likely to be any integer between 21 and 120 (inclusive). Use Bayes’ Theorem (and software) to determine the posterior probabilities, conditional on the observed data. Produce a graph of these posterior probabilities as a function of $N$. Given the data, which value(s) of $N$ is/are most likely, and which is/are least likely?

Applying Bayes’ Theorem in the same way but with these new prior probabilities (.01 for the 100 possible values between 21 and 120, inclusive) gives the following graph of posterior probabilities:
The most likely values of $N$ are $N=49$ and $N=50$, which each have posterior probability .0198.

The least likely values of $N$ are $N=21, 22, 23, 24, 25$, all of which have posterior probability 0.

*Not to turn in:* Consider working on exercises 1-15 at the end of section 2.3 in the text.