1. [3 points] What are the axioms of probability?

Solution:
- **Normality**: $0 \leq P(X) \leq 1$
- **Certainty**: $P(\Omega) = 1$
- **Additivity**: $P(X \lor Y) = P(X) + P(Y)$, when $X$ and $Y$ are mutually exclusive; or
  - **Additivity**: $P(X \lor Y) = P(X) + P(Y) - P(X \& Y)$ (general/overlap form)

2. [2 points] (a) What two definitions of probability supplement these axioms?

Solution:
- **Independence**: $P(X \& Y) = P(X) \cdot P(Y)$, when $X$ and $Y$ are independent.
- **Conditional Probability**: $P(X|Y) = \frac{P(X\&Y)}{P(Y)}$

(b) [1 point] What is Bayes’ Rule?

Solution: $P(X|Y) = \frac{P(X) \cdot P(Y|X)}{P(X) \cdot P(Y|X) + P(\sim X) \cdot P(Y|\sim X)}$

or

$P(X|Y) = \frac{P(X) \cdot P(Y|X)}{P(Y)}$

3. [2 points] What is a valid argument?

Solution: An argument is valid if there are no counterexamples, where a counterexample is a case of all true premises leading to a false conclusion. (Other definitions of validity are also acceptable.)

4. [2 points] What does it mean for two events (or propositions) to be mutually exclusive? To be independent?

Solution:
- **Mutually Exclusive**: Two events are mutually exclusive if they both cannot occur, i.e., if the occurrence of one precludes the occurrence of the other;
  - Two propositions are mutually exclusive if they both cannot be true at the same time, i.e., the truth of one entails the falsity of the other.
- **Independence**: Two *events* are independent when the occurrence of one has no bearing on the probability of occurrence of the other;
  Two *propositions* are true when the truth of one has no bearing on the probability of the truth of the other.
5. Two events, A and B, are such that $P(A)=0.4$, $P(B)=0.6$, and $P(A \& B)=0.3$

(a) [2 points] Are A and B independent? Provide an argument or proof to support your answer.

**Solution:** No, A and B are not independent. If they were, $P(A \& B)$ would equal 0.24, but it does not.

(b) [2 points] Are A and B mutually exclusive? Provide an argument or proof to support your answer.

**Solution:** No, A and B are not mutually exclusive. If they were, $P(A \& B)$ would equal 0, but it does not.

(c) [2 points] Are A and B exhaustive? Provide an argument or proof to support your answer.

**Solution:** No, A and B are not exhaustive. If they were, $P(A \lor B)$ would = 1, but it does not. $P(A \lor B) = 0.4 + 0.6 - 0.3 = 0.7$

Solve for the following:

(d) [2 points] $P(A \lor B)$

**Solution:** $P(A \lor B) = 0.4 + 0.6 - 0.3 = 0.7$

(e) [2 points] $P(B|A)$

**Solution:** $P(B|A) = \frac{P(B \& A)}{P(A)} = \frac{0.3}{0.4} = \frac{3}{4} = 0.75$

(f) [2 points] $P(\sim A)$

**Solution:** $P(\sim A) = 1 - P(A) = 1 - 0.4 = 0.6$

Or use a Venn diagram to work it out.

(g) [2 points] $P(\sim B)$

**Solution:** $P(\sim B) = 1 - P(B) = 1 - 0.6 = 0.4$

Or use a Venn diagram to work it out.

(h) [2 points] $P(\sim A \& \sim B)$

**Hint:** Try using a Venn diagram to figure this out.

**Solution:** $P(\sim A \& \sim B) = 0.3$
6. Being a connoisseur of donuts, Prof. Haber opens a donut shop, “Inductive Donuts.” Being a probability enthusiast, Prof. Haber decides that he will not reveal the donut menu, but instead require people to infer the donut varieties from random samplings of donuts. For every donut taken, Prof. Haber replaces it with the same variety.

Upon some rather delicious detective work, you’ve determined that 50% of the donuts in Prof. Haber’s shop are made with chocolate. You’ve also been able to determine that the probability that a donut will be frosted, given that it is chocolate, is 30%.

(a) What is the probability that a donut is both frosted and chocolate?

\[
P(F \cap C) = \frac{P(F \cap C)}{P(C)}
\]

\[
0.3 = \frac{P(F \cap C)}{0.5}
\]

so, \( P(F \cap C) = 0.15 \)

(b) Prof. Haber tells you that a donut being frosted or being made with chocolate are independent events. Given what you’ve already inferred, what is the probability that a donut in Prof. Haber’s shop is frosted?

Solution: Since \( F \) and \( C \) are independent, \( P(F \cap C) = P(F) \cdot P(C) \).

Since \( P(F \cap C) = 0.15 \), and \( P(C) = 0.5 \), \( P(F) \) must equal 0.3 if they are independent.

(c) As you continue to sample Prof. Haber’s truly lovely donuts, you discover that whether a donut is glazed has no bearing on the probability of that donut having sprinkles (and vice-versa). Additionally, some donuts are neither glazed nor sprinkled.

Assign a probability to a donut being glazed, and to a donut having sprinkles, that reflect these two facts. Prove that your probabilities reflect that these are independent but not exhaustive of the donuts.

Solution: Because \( G \) and \( S \) are independent, \( P(G \cap S) \) will equal \( P(G) \cdot P(S) \), but \( P(G \lor S) < 1 \).

(d) Every Inductive Donut is either baked or fried. No donut is both baked and fried, because Prof. Haber thinks that is an abomination. Assign probabilities that reflect these facts, and offer a proof of that.

Solution: Because \( B \) and \( F \) are exclusive, \( P(B \cap F) \) must equal 0. Because they are exhaustive, \( P(B \lor F) \) must equal 1.
7. Scandal hits Inductive Donuts! A diligent local reporter (perhaps a disgruntled former student?) discovers and reveals that some of the donuts are not, as claimed, made in shop, but out-sourced. They are fakes! 10% of jelly donuts, 20% of custard and (most shockingly) 80% of donut holes are not made in-house.

The reporter also reveals that of the donuts at Inductive Donuts, 20% are jelly, 10% custard, and 10% donut holes. These are exclusive of each other. No other donuts are involved in the scandal (i.e., all other donuts are made at the shop).

(a) [6 points] Your friend finds out the donut they just ate was not made in-house. Given that information, what is the probability your friend ate a donut hole? Demonstrate this using both Bayes’ Rule and the Probability Tree method.

**Solution: Bayes’ Rule:**

What we know:
key: F = Fake donut; J = Jelly; C = Custard; H = Donut Hole
\[ P(F|J) = 0.1; \quad P(J) = 0.2 \]
\[ P(F|C) = 0.2; \quad P(C) = 0.1 \]
\[ P(F|H) = 0.8; \quad P(H) = 0.1 \]

Now solve for \( P(H|F) \):

\[
P(H|F) = \frac{P(H) \cdot P(F|H)}{P(H) \cdot P(F|H) + P(C) \cdot P(F|C) + P(J) \cdot P(F|J)}
\]
\[
= \frac{0.1 \cdot 0.8}{0.1 \cdot 0.8 + 0.1 \cdot 0.2 + 0.2 \cdot 0.1}
\]
\[
= \frac{0.08}{0.08 + 0.02 + 0.02} = \frac{0.08}{0.12} = \frac{2}{3} = 0.6667 = 66.7\%
\]

For the probability tree, there are a few different options. Your first split might be a four-way split between H (donut holes), C (custard), J (jelly) and O (other), with each branch receiving a probability value of 0.1, 0.1, 0.2, and 0.6, respectively. You can disregard the ‘Other’ branch, since from the information you know that \( P(F|O) = 0 \). The second set of branches will be ‘F’ (Fake) or ‘\~{}F’ (not Fake), for each of the three kinds of donuts. Then solve for \( P(H|F) \) using the definition of conditional probability and the information from the tree.

(b) [6 points] Your friend can tell, with 70% accuracy, whether a custard donut is genuine or fake. Prof. Haber has said that if you can demonstrate that the probability that you’ve been given a fake donut is greater than 50%, that you will receive a refund.

You get a custard donut, which your friend says is a fake. You eat it anyways, then get another custard donut, which your friend again says is fake. What probability should you assign that you’ve received a fake donut? Is it enough to get a refund?

**Solution:** This is similar to Hacking’s taxicab problem, and Haber’s snipe problem, so let’s use Bayes’ Rule:
What we know:

key: \( R_F \) = Friend reports that a donut is fake.
\[ P(R_F|F) = 0.7 \quad P(R_F|\sim F) = 0.3 \]

First solve for \( P(F|R_F) \):

\[
P(F|R_F) = \frac{P(F) \cdot P(R_F|F)}{P(F) \cdot P(R_F|F) + P(\sim F) \cdot (P(R_F|\sim F))}
\]
\[
= \frac{0.2 \cdot 0.7}{0.2 \cdot 0.7 + 0.8 \cdot 0.3}
\]
\[
= \frac{0.14}{0.14 + 0.24} = \frac{14}{38} = 0.368 = 36.8\%
\]

We can now use \( \frac{7}{19} \) as our new prior probability that we had a fake donut when we run Bayes’ a second time (for the second custard donut). Your friend’s reliability of reporting does not change, i.e., they are still correct 70% (\( \frac{7}{10} \)) of the time.

\[
P(F|R_F) = \frac{P(F) \cdot P(R_F|F)}{P(F) \cdot P(R_F|F) + P(\sim F) \cdot (P(R_F|\sim F))}
\]
\[
= \frac{\frac{7}{19} \cdot \frac{7}{10}}{\frac{7}{19} \cdot \frac{7}{10} + \frac{12}{19} \cdot \frac{3}{10}}
\]
\[
= \frac{49}{190} + \frac{36}{190} = \frac{49}{190} = \frac{49}{85} = 0.576 = 57.6\%
\]

So you can demonstrate to Prof. Haber that given your friend’s accuracy of reporting, that you have assigned a greater than 50% probability to having received a fake donut. Thus you get your refund!
8. One of Ian Hacking’s ‘Odd Questions’ concerns witnessing taxicabs. Here is the set up:

There are two kinds of taxis in town, Green and Blue. Green cabs dominate the market, with 85% of the cabs on the road. Like the Blue cabs, they are randomly distributed about town.

On a misty winter night, a taxi sideswiped another car and drove off. A witness says it was a blue cab. That witness is tested, and 80% of the time she correctly reports the color of the cab in conditions like those on the night of the accident. She was just as accurate in reporting a blue or green cab. Hacking demonstrated what probability we might assign to the cab being blue, given the report of the witness.

Here’s the twist. You are on the jury, serving with other members of this class. The judge instructs you that may not conclude that the cab was one color or the other with confidence unless you can assign a probability of greater than 65% to that belief. You’ve heard from the first witness, and the prosecution says they have several more witnesses lined up can report, with the same level of accuracy, what color the cab was.

Starting from the data provided, how many witnesses will you need to see before you can assign a probability of greater than 65% to the cab being blue? Prove this using both Bayes’ Rule and a Probability Tree.

**Solution:** Using Bayes’ rule:

First, list what we know:

- \( P(G) = 85\% \); \( P(W_B|G) = 20\% \)
- \( P(B) = 15\% \); \( P(W_B|B) = 80\% \)

What we are looking for is how many witnesses we will need to hear from before we can assign \( P(B) > 65\% \). To figure that out, we’ll need to determine the probability that the cab is blue upon each subsequent witness testifying as such, identifying the \( x^{th} \) witness that pushes that probability over 65%, i.e., we need \( P(B|W_x B) \).

So let’s start things off, and keep running Bayes’ until we get \( P(B|W_x B) > 65\% \):

\[
P(B|W_B) = \frac{P(B) \cdot P(W_B|B)}{P(B) \cdot P(W_B|B) + P(G) \cdot P(W_B|G)}
\]

\[
= \frac{0.15 \cdot 0.8}{0.15 \cdot 0.8 + 0.85 \cdot 0.2} = \frac{0.12}{0.12 + 0.17} = \frac{0.12}{0.29} = 0.4137 \approx 41.4\%
\]

41.4% is less than 65%, so let’s see what happens when the second witness testifies and we use \( \frac{12}{29} \) as our new prior probability:

\[
P(B|W_B) = \frac{P(B) \cdot P(W_B|B)}{P(B) \cdot P(W_B|B) + P(G) \cdot P(W_B|G)}
\]

\[
= \frac{\frac{12}{29} \cdot \frac{8}{10}}{\frac{12}{29} \cdot \frac{8}{10} + \frac{17}{29} \cdot \frac{2}{10}} \quad (\text{You might choose not to use fractions})
\]

\[
= \frac{96}{290} + \frac{34}{290} = \frac{96}{130} \approx 73.9\%
\]

So after two witnesses, \( P(B) \) will be greater than 65%.
A medical diagnostic test correctly reports the presence (or absence) of disease 99% of the time. The overall rate of disease in the population is $\frac{1}{10,000}$.

What probability should a doctor assign that someone who tested positive has the disease? If the doctor runs the test again, and it comes out positive, what probability should she assign to that person having the disease?

Show your results using both Bayes’ Rule and the Probability Tree method.

**Solution:** I find this problem easier to do using a probability tree, but those are more difficult to typeset, so let’s use Bayes’ Rule here.

First let’s list what we know:

- $P(D) = \frac{1}{10,000} = 0.01\%$; $P(+|D) = \frac{99}{100} = 99\%$
- $P(\sim D) = \frac{9999}{10,000} = 99.99\%$; $P(+|\sim D) = \frac{1}{100} = 1\%$

key: D = Has disease; + = Tests positive for disease

Now let’s solve for $P(D|+)$:

\[
P(D|+) = \frac{P(D) \cdot P(+|D)}{P(D) \cdot P(+|D) + P(\sim D) \cdot P(+|\sim D)}
\]

\[
= \frac{1/10,000 \cdot 99/100}{1/10,000 \cdot 99/100 + 9999/10,000 \cdot 1/100}
\]

\[
= \frac{99}{1,000,000} + \frac{9999}{1,000,000} = \frac{99,10098}{1,000,000} = \frac{99}{10098} \approx 0.0098 \approx 0.98\%
\]

Pretty low! Let’s see what happens if the patient tests positive a second time, using 0.98% as our new prior:

\[
P(D|+) = \frac{P(D) \cdot P(+|D)}{P(D) \cdot P(+|D) + P(\sim D) \cdot P(+|\sim D)}
\]

\[
= \frac{0.98\% \cdot 99\%}{0.98\% \cdot 99\% + 99.02\% \cdot 1\%}
\]

\[
= \frac{0.9702\%}{0.9702\% + 0.9902\%} \approx 0.4949 \approx 49.5\%
\]

Big difference! Please note that you can calculate this using fractions, decimals, or percentages.

For the probability tree, your first split should be $D$ or $\sim D$, then + or $\sim+$ for each. This displays nicely how many more false positives there are than true positives in the population.
10. Last night the St. Louis Cardinals (STL) defeated the Texas Rangers (TEX) in game 1 of Major League Baseball’s World Series. The first team to win four games wins the Series, and is crowned champion for the season. The teams are evenly matched, but playing at home gives each team a slight advantage. As a result, the probability of winning a home game is 60%. Game 2 will be played in St. Louis, games 3, 4, and 5 in Texas, and games 6 and 7 will be back in St. Louis.

(a) Assume that the St. Louis Cardinals win game 4. What is the probability of that win being the fourth win for STL in the Series? That is, given that STL wins game 4, what is the probability they swept the Series and are champions? 
(\textit{Note: Remember that St. Louis has already won game 1!})

\textbf{Solution:} This is easily done with a probability tree. After all, there are many ways in which STL might win game 4, but only one of those is as part of a four game sweep.

Key: \( W_x \) = Win game \( x \); \( L_x \) = Lose game \( x \)

What we know:
\[
\begin{align*}
P(W_1) &= 1.0; & P(L_1) &= 0.0 & \text{(They already won the game)} \\
P(W_2) &= 0.6; & P(L_2) &= 0.4 \\
P(W_3) &= 0.4; & P(L_3) &= 0.6 \\
P(W_4) &= 0.4; & P(L_4) &= 0.6
\end{align*}
\]

Goal: Solve for \( P(W_2 \& W_3 | W_1 \& W_4) \)

Now we can easily solve for \( P(W_2 \& W_3 | W_1 \& W_4) \) using the definition of conditional probability. Note that there are four ways the Cardinals might win game 4, given that they’ve already won game 1. Let’s label these a, b, c, and d, corresponding from left to right in the tree above. Since they are on distinct branches, they are mutually exclusive of each other.

So \( P(W_4) = P(W_{4a} \lor W_{4b} \lor W_{4c} \lor W_{4d}) = 0.096 + 0.144 + 0.064 + 0.096 = 0.4 \).

Furthermore, the probability of winning a game is independent of winning or losing a previous game (from setup). This, along with the definition of conditional probability
give us:

\[
P(W_2 \& W_3|W_1 \& W_4) = \frac{P(W_1 \& W_2 \& W_3 \& W_4)}{P(W_1 \& W_4)}
\]

\[
= \frac{1.0 \cdot 0.6 \cdot 0.4 \cdot 0.4}{1.0 \cdot 0.4}
\]

\[
= \frac{0.96}{0.4} = 0.24 = 24\%
\]

**Note:** Since \(P(W_1) = 1.0\), you can effectively drop it out. I’ve left it in here so you can see the details when you leave it in.

**Note:**
You might also notice that winning games are independent events. In that case, you can simply use the definition of conditional probability to work out the answer. The probability tree will help organize that information, but isn’t necessary.

(b) If the Texas Rangers end up winning the world series, what is the probability that they won their fourth game of the Series in game 6 of the World Series? That is, given that the Texas Rangers won the world series, what is the probability that they did so by winning four of the next five games?

**Solution:** Given that the Texas Rangers lost game 1, the subsequent event “Texas Rangers win the World Series in 6 games” (call this \(T_6\)) describes several different parts of the probability (or sample) space, namely the following sequences of W(ins) and L(osses):

\((LWWW)\); \((WLWWW)\); \((WWLWW)\); & \((WWWLW)\)

So, \(T_6\) is just a description of these four separate sequences of wins/losses. Notice that they are mutually exclusive of each other, and that we can describe this formally as:

\(T_6 = LWWW \vee WLWWW \vee WWLWW \vee WWWLW\)

(We haven’t proved it yet, but this is an instance of Total Probability)

By the axiom of additivity, we can easily work out \(P(T_6)\):

\[
P(T_6) = P(LWWW \vee WLWWW \vee WWLWW \vee WWWLW)
\]

\[
= P(LWWW) + P(WLWWW) + P(WWLWW) + P(WWWLW)
\]

\[
= (0.6 \cdot 0.6 \cdot 0.6 \cdot 0.6 \cdot 0.4) + (0.4 \cdot 0.4 \cdot 0.6 \cdot 0.6 \cdot 0.4)
\]

\[
+ (0.4 \cdot 0.6 \cdot 0.4 \cdot 0.6 \cdot 0.4) + (0.4 \cdot 0.6 \cdot 0.6 \cdot 0.4 \cdot 0.4)
\]

\[
= 0.05814 + 0.02304 + 0.02304 + 0.02304 = 0.12726 \approx 12.7\%
\]
11. Abigail, Beauregard, Chisolm, and Darlene are finalists on *The Worlds Worst Reality TV Show*. *The Worlds Worst Reality TV Show* has three remaining possible challenges. Which challenge comes up is determined by the drawing of straws (see below). Abigail won the previous challenge, so she has an advantage going into the final rounds: she gets to draw straws first to determine which challenge she will play in the first round, and can then select her opponent for that round. The other two remaining challengers will face off, and the winners of each challenge will then play in the championship round.

In other words, if Abigail wins the first round, she will then face one of the other two competitors not faced before. The championship round challenge will be an invigorating game of ‘wff ’n proof’.

Below is a table with the following information. It tells you the probability of any one challenge being selected for the first round, and the probability of Abigail beating her competitors in each challenge. It also provides you with information about the final challenge.

<table>
<thead>
<tr>
<th></th>
<th>American Idol Outtakes</th>
<th>I ate WHAT?!</th>
<th>Win Ayn Rand’s Money!</th>
<th>wff ’n proof (final round)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability of Being Selected for first round</td>
<td>30%</td>
<td>50%</td>
<td>20%</td>
<td>NA</td>
</tr>
<tr>
<td>Beauregard</td>
<td>60%</td>
<td>70%</td>
<td>40%</td>
<td>60%</td>
</tr>
<tr>
<td>Chisolm</td>
<td>40%</td>
<td>60%</td>
<td>50%</td>
<td>75%</td>
</tr>
<tr>
<td>Darlene</td>
<td>70%</td>
<td>50%</td>
<td>30%</td>
<td>25%</td>
</tr>
</tbody>
</table>

Finally, the probability of Beauregard beating Chisolm in any challenge is 60%, and of beating Darlene is 30%. Darlene has a 50% probability of beating Chisolm in any challenge.

(a) What is the probability that Abigail will win *The Worlds Worst Reality TV Show*?

(b) If Abigail wins *The Worlds Worst Reality TV Show*, what is the probability that she faced Beauregard in the final rounds?

(c) Assume Abigail faced Darlene in the first round. What is the probability that Darlene faced Chisolm in the final round if Darlene ultimately won *The Worlds Worst Reality TV Show*?

(Hint: If you use %’s in your calculations, rounding up to whole numbers is allowed.)
12. Given the axioms of probability, the definition of independence, conditional probability, and the rules of logic and algebra, demonstrate how the following version of Bayes’ Rule may be derived:

\[ P(X|Y) = \frac{P(X) \cdot P(Y|X)}{P(X) \cdot P(Y|X) + P(\neg X) \cdot P(Y|\neg X)} \]

**Solution:**

Start with Definition of Conditional Probability:

\[ P(X|Y) = \frac{P(X \& Y)}{P(Y)} \]

From this we can solve for \( P(X \& Y) \)...

\[ P(Y) \cdot P(X|Y) = \frac{P(X \& Y)}{P(Y)} \cdot P(Y) \]

\[ P(X \& Y) = P(X|Y) \cdot P(Y) \]

Recall that \( P(X \& Y) = P(Y \& X) \) so

\[ P(Y|X) = \frac{P(Y \& X)}{P(X)} \]

\[ P(X) \cdot P(Y|X) = \frac{P(Y \& X)}{P(X)} \cdot P(X) \]

\[ P(X \& Y) = P(Y \& X) = P(Y|X) \cdot P(X) \]

From logic, the law of total probability:

\[ P(Y) = P(Y \& X) + P(Y \& \neg X) \]

(Hint: what is \( Y \) equivalent to on an X-Y truth table? On a Venn diagram?)

Substitute this for \( P(Y) \) in the definition of conditional probability:

\[ P(X|Y) = \frac{P(X \& Y)}{P(Y \& X) + P(Y \& \neg X)} \]

Now substitute for \( P(X \& Y) \) and \( P(X \& \neg Y) \), using the \( P(X) \) formulation:

\[ P(X|Y) = \frac{P(Y|X) \cdot P(X)}{[P(Y|X) \cdot P(X)] + [P(Y \& \neg X) \cdot P(\neg X)]} \]

Now simply rearrange to get the desired formulation of Bayes’ Rule:

\[ P(X|Y) = \frac{P(X) \cdot P(Y|X)}{P(X) \cdot P(Y|X) + P(\neg X) \cdot P(Y|\neg X)} \]
<table>
<thead>
<tr>
<th>Question:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Points:</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>16</td>
<td>12</td>
<td>12</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>90</td>
</tr>
<tr>
<td>Score:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>