## The Three Probability Axioms

PA1 **Nonnegativity**: For any statement $A$, $\Pr(A) \geq 0$.

PA2 **Certainty**: If you're certain that a statement $A$ is true, then $\Pr(A) = 1$ (and if you're certain that $A$ is false, then $\Pr(A) = 0$).

PA3 **Additivity** If $A$ and $B$ are never true at the same time, then $\Pr(A \vee B) = \Pr(A) + \Pr(B)$.

The following is a consequence of the probability axioms:

1. **Partitionality**: If $H_1, H_2, \ldots H_n$ are statements and you know that *exactly one* of them is true, then $\Pr(H_1) + \Pr(H_2) + \cdots + \Pr(H_n) = 1$.

## The Ratio Formula and Theorems based on it

1. **The Ratio Formula for Conditional Probability**: If $A$ and $B$ are statements and $\Pr(B) \neq 0$, then $\Pr(A \mid B) = \frac{\Pr(A \& B)}{\Pr(B)}$.
2. **Bayes' Theorem**: If $E$ and $H$ are statements, then

$$\Pr(H \mid E) = \Pr(H)\frac{\Pr(E \mid H)}{\Pr(E)}.$$

3. **Total Probability**: If $H_1, H_2, \ldots H_n$ are statements and you know that *exactly one* of them is true, then for any statement $E$:

$$\Pr(E) = \Pr(E \mid H_1)\Pr(H_1) + \Pr(E \mid H_2)\Pr(H_2) \cdots + \Pr(E \mid H_n)\Pr(H_n).$$

If there exactly two hypotheses $H$ and $\neg H$, then

$$\Pr(E) = \Pr(E \mid H)\Pr(H) + \Pr(E \mid \neg H)\Pr(\neg H).$$

## Conditionalization

**The Conditionalization Norm**: If you have confidence $\Pr(H)$ in a hypothesis $H$, and then you learn $E$, then your new confidence in $H$ upon learning $E$, $Pr_E(H)$, should equal your conditional probability in $H$ given $E$:

$$Pr_E(H) \overset{\text{should}}{=} Pr(H \mid E).$$

## Evidence For and Independence

1. **Definition**:
   (a) $E$ is **evidence for** $H$ if $\Pr(H \mid E) > \Pr(H)$.
   (b) $E$ is **evidence against** $H$ if $\Pr(H \mid E) < \Pr(H)$.
   (c) $H$ is **independent of** $E$ if $\Pr(H \mid E) = \Pr(H)$.
2. **The Evidence-For Lemma**: The following four are equivalent:

$$\Pr(H \mid E) > \Pr(H) \qquad \Pr(E \mid H) > \Pr(E)$$

$$\Pr(H \mid E) > \Pr(H \mid \neg E) \qquad \Pr(E \mid H) > \Pr(E \mid \neg H).$$

They are still equivalent if the ">" are replaced with "<" or "="

## 1 | Approaching Problems on Midterm 2

The most important things to know how to use, and when to use, are **Bayes' Theorem**, **Total Probability**, and the **Evidence-For Lemma**. Bayes' Theorem and Total Probability are trickier, so here is how I recommend going about the problems on this midterm:

**Setting Up your Problems**:

1. The problem will give you some setup. Even before you try to answer the question, *write down* all the probabilities, both unconditional (with no middle bar) and conditional (with a middle bar), that you already know given the setup of the problem. Give the different statements letter names for convenience, like $B$ for "we choose the blue bag" or $F$ for "Freya will be here."
2. For any pair of statements where exactly one of them is true (i.e. if you know that a coin will either come up heads or tails), if you know the probability of one, write down the probability of the other by subtracting it from 1. For instance, if we have a biased coin that comes up heads 0.7 of the time, $Pr(\text{heads}) = 0.7$, then $Pr(\text{tails}) = 1 - Pr(\text{heads}) = 0.3$.
3. Write out exactly what the problem statement is asking for, in a form like $Pr(A \mid B)$ or $Pr(C)$.
4. If a problem statement says that you learn something (e.g., you flip a coin and it lands heads) and then asks you for your new opinion, like this diagram:
   $$Pr(H) \longrightarrow \text{learn } E \longrightarrow Pr_E(H)$$
   In this circumstance, you should assume that you conditionalize. If $Pr_E$ is your new opinion upon learning $E$, **conditionalization** says you should have
   $$Pr_E(H) = Pr(H \mid E).$$
   So a problem that asks you for your new opinion about $H$ after you learn $E$ is likely asking you for $Pr(H \mid E)$.
5. Solve the problem. At each intermediate step, write out in symbols what you're doing, so if you get the numbers wrong you can go back and understand what you meant to do.

**Solving in General:** In general, the problem statement will either ask you for a *conditional* probability like $Pr(A \mid B)$, with a bar in the middle, or an *unconditional* probability like $Pr(A)$, with no bar in the middle. If there is a bar in the middle, you should probably use **Bayes' Theorem**. If there is *no* bar in the middle, you should probably use **Total Probability**. Below I explain the strategy for each.

**Solving – Bayes' Theorem**:

1. In general, if the problem is asking you for something like $Pr(A \mid B)$ with a bar in the middle (i.e. a conditional probability), *most likely* you will have to use Bayes' Theorem:
   $$Pr(A \mid B) = Pr(A)\frac{Pr(B \mid A)}{Pr(B)}.$$

2. Notice that Bayes' Theorem "flips" a conditional probability: If you know $Pr(B \mid A)$ then you can find $Pr(A \mid B)$.
3. Most likely, if you're being asked for $Pr(A \mid B)$, then the problem statement already gives you $Pr(B \mid A)$ and $Pr(A)$. So you can plug these into the formula.
4. But often you won't know $Pr(B)$. What do you do then? In most cases, you can probably calculate $Pr(B)$ using **Total Probability**: see below.
5. Once you have all the numbers, do the math to get the final answer.

**Solving – Total Probability**:

1. Suppose you are trying to figure out some *unconditional* probability like $Pr(B)$. If the problem statement gives you probabilities like $Pr(B \mid A_1)$, $Pr(B \mid A_2)$, $Pr(A_1)$, and $Pr(A_2)$, and tells you that exactly one of $A_1$ and $A_2$ (or there may be more than two) is true (e.g., the bag is either blue or red), then you can use total probability to solve for $Pr(B)$:

$$Pr(B) = Pr(B \mid A_1)Pr(A_1) + Pr(B \mid A_2)Pr(A_2)$$

   where you can keep adding terms $Pr(B \mid A_i)Pr(A_i)$ if there are more hypotheses and you know that *exactly* one of them is true.
2. Often, in the middle of a calculation of Bayes' theorem, you will need to use total probability to get a $Pr(B)$.

## 2 | Chocolates

Here I'll solve the "Chocolates" problem from the midterm review session using the method outlined above.

1. Willy Wonka Chocolates Inc. makes two kinds of boxes of chocolates. The "wonk box" has four caramel chocolates and six regular chocolates. The "zonk box" has six caramel chocolates, two regular chocolates, and two mint chocolates. A third of their boxes are wonk boxes, the rest are zonk boxes.
2. They don't mark the boxes. The only way to tell what kind of box you've bought is by trying the chocolates inside. In fact, all the chocolates look the same; you can only tell the difference by tasting them.

→ If you buy a random box, try a chocolate at random, and find that it's caramel, what is the probability you've bought a wonk box? (Weisberg, *Odds and Ends*)

**Solution**: First, let's write down all the things we already know. Let $W$ be the statement "you've bought a wonk box" and $Z$ be the statement "you've bought a zonk box". Let $C$ be the statement "you tried a caramel chocolate", $R$ be the statement "you tried a regular chocolate", and $M$ be the statement "you tried a mint chocolate".

**What is the Problem Asking?** The problem statement says that we've tried a chocolate, and it turns out to be caramel. So you've learned the statement $C$. The problem statement then asks what the new probability is that you've bought a wonk box. That is, it is asking for $Pr_C(W)$. Given this, let's use **Conditionalization**. Conditionalization says $Pr_C(W)$ should equal $Pr(W \mid C)$. So really the problem statement is asking us for $Pr(W \mid C)$. Let's solve this:

1. **Setting Up**: Given the setup of the case, we already know the following probabilities:

$$Pr(W) = 1/3 \qquad Pr(Z) = 2/3$$

$$\begin{aligned} Pr(C \mid W) &= 0.4 & Pr(C \mid Z) &= 0.6 \\ Pr(R \mid W) &= 0.6 & Pr(R \mid Z) &= 0.2 \\ Pr(M \mid W) &= 0.0 & Pr(M \mid Z) &= 0.2 \end{aligned}$$

   where $Pr(Z) = 1 - Pr(W) = 1 - 1/3 = 2/3$.

2. **Finding** $Pr(W \mid C)$: Should we use Bayes' Theorem or Total Probability to figure out $Pr(W \mid C)$? Because this is a *conditional* probability, with a bar in the middle, we should use **Bayes' Theorem**.

   We write out Bayes' Theorem with the right statement letters:

$$Pr(W \mid C) = Pr(W)\frac{Pr(C \mid W)}{Pr(C)}.$$

   Now, we already know (from our list above) that $Pr(W) = 1/3$ and $Pr(C \mid W) = 0.4$. So we can substitute these in:

$$Pr(W \mid C) = 1/3\frac{0.4}{Pr(C)}.$$

   But we don't know what $Pr(C)$ is, so we'll have to calculate it.
3. **Finding** $Pr(W \mid C)$: Should we use Bayes' Theorem or Total Probability to figure out $Pr(W \mid C)$? Because this is a *unconditional* probability, with *no* bar in the middle, we should use **Total Probability**.

   What are the two ways in which $C$ could be true? Well, we could either have gotten a caramel out of a wonk box, or gotten a caramel out of the zonk box. So the two hypotheses, exactly one of which is true, are $W$ and $Z$. Let's write out total probability with the right statement letters:

$$Pr(C) = Pr(C \mid W)Pr(W) + Pr(C \mid Z)Pr(Z).$$

   And we see from above that we know all these values already, so let's substitute them in:

$$Pr(C) = 0.4 \cdot 1/3 + 0.6 \cdot 2/3 = 8/15.$$

4. **Finding** $Pr(W \mid C)$**, resumed**: Now that we have $Pr(C) = 8/15$ we can substitute it back into Bayes' Theorem:

$$Pr(W \mid C) = 1/3\frac{0.4}{8/15} = 1/4.$$

So our final answer is $Pr(W \mid C) = 1/4$.