

Unit 1 Homework

w203: Statistics for Data Science; Thursdays, Section 6: Samantha Williams

This assignment is due at the time of your live session. Notice that there are two parts that are mandatory: a section called **Applied Practice**, and a section called **Proof Practice**. There is also an optional programming assignment, if you're interested.

Section 1: Applied Practice

For each of these questions, please do the work (probably with a pen and paper), and then enter your solutions into the Gradescope assignment slot that is called, "Unit 1 – Applied Practice."

1. (9 points total) Summer Roadtrip

At gas station on Route 66, 40% of customers use regular gas (event R), 35% use mid-grade gas (event M), and 25% use premium gas (event P). Of the customers that use regular gas, 30% fill their tanks (event F). Of the customers that use mid-grade gas, 60% fill their tanks, while of those that use premium, 50% fill their tanks. Assume that each customer is drawn independently from the entire pool of customers.

WE KNOW: Regular Gas $P(R) = 0.4$

Mid-grade Gas $P(M) = 0.35$

Premium Gas $P(P) = 0.25$

Fill Tank given Regular Gas $P(F|R) = 0.3$

Fill Tank given Mid-grade Gas $P(F|M) = .6$

Fill Tank given Premium Gas $P(F|P) = .5$

1. (3 points) What is the probability that the next customer will request regular gas and fill the tank?

Using Conditional Probability $P(A|B) = \frac{P(A \cap B)}{P(B)} \Rightarrow$ rearranged becomes the Multiplicative Law

$$P(A \cap B) = P(A|B)P(B)$$

$$P(R \cap F) = P(F|R)P(R)$$

$$P(R \cap F) = 0.3 \times 0.4 = 0.12$$

2. (3 points) What is the probability that the next customer will fill the tank with any kind of gas?

Using Total Probability $P(B) = \sum_i P(B \cap A_i)$ or it can be written as $P(B) = \sum_i P(B|A_i)P(A_i)$

$$P(F) = P(F|R)P(R) + P(F|M)P(M) + P(F|P)P(P)$$

$$P(F) = (0.3 \times 0.4) + (0.6 \times 0.35) + (0.5 \times 0.25) = 0.12 + 0.21 + 0.125 = 0.455$$

3. (3 points) Given that the next customer fills the tank, what is the conditional probability that they use regular gas?

Using Bayes' Rule $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$

$$P(R|F) = \frac{P(F|R)P(R)}{P(F)}$$

$$P(R|F) = \frac{P(0.3)(0.4)}{P(0.455)} = 0.264$$

2. (12 points total) The Claaaaaw

Suppose that in a claw game at an arcade, there is a collection of toys that have the following characteristics:

- $2/5$ are red;
- $3/5$ are waterproof;
- $1/2$ are cool.

(When we write that $2/5$ are red, this means that $3/5$ are not red. But the facts above do not provide any information about the intersection between the different traits.)

Furthermore:

- $1/5$ are both red *and* waterproof;
- $1/5$ are both red *and* cool;
- $3/10$ are both waterproof *and* cool.

(When we write that $1/5$ are both red *and* waterproof, this contains no information about how cool they are.)

Finally:

- $1/10$ are neither red, waterproof, nor cool. (These are pretty lame toys.)

Since working those claws is so hard, suppose that any toy in the game has an equal chance of being selected.

WE KNOW: Red $P(R) = \frac{4}{10}$

Waterproof $P(W) = \frac{6}{10}$

Cool $P(C) = \frac{5}{10}$

Red and Waterproof $P(R \cap W) = \frac{2}{10}$

Red and Cool $P(R \cap C) = \frac{2}{10}$

Waterproof and Cool $P(W \cap C) = \frac{3}{10}$

Based on Kolmogorov Axioms: Unitarity $P(\Omega) = 1$ and Non-negativity $\forall A \in S, P(A) \geq 0$ where $P(A)$ is finite and Real. Apply a basic property of probability, the complement rule $P(A^C) = 1 - P(A)$:

$P(R^C) = 1 - \frac{4}{10} = \frac{6}{10}$ are not red

$P(W^C) = 1 - \frac{6}{10} = \frac{4}{10}$ are not waterproof

$P(C^C) = 1 - \frac{5}{10} = \frac{5}{10}$ are not cool

Given that $P(R \cap W \cap C)^C = \frac{1}{10}$ are not red, waterproof and cool we know that $P(R \cup W \cup C) = \frac{9}{10}$ are red, waterproof or cool or all three.

1. (3 points) Draw an area diagram to represent these events. For as many of the events that you can, compute the probability (for example, $P(R \cap C \cap W)$)).

![Area Diagram](~/Images/area_diagram.png) ### Could not knit with it on document see other page

2. (3 points) What is the probability of drawing a toy that is red *and* waterproof *and* cool?

$$P(R \cap W \cap C) = \frac{1}{10}$$

3. (3 points) Suppose that you pull out a toy at random, and you observe only the color, noting that it is red. Conditional on just this information, what is the probability that the toy is not cool?

$$P(C^c|R) = \frac{P(C^c \cap R)}{P(R)} = \frac{2/10}{4/10} = \frac{1}{2}$$

4. (3 points) Given that a randomly selected toy is *either* red or waterproof, what is the probability that it is cool?

$$P(C|R) = \frac{P(C \cap R)}{P(R)} = \frac{\frac{2}{10}}{\frac{4}{10}} = 1/2 \quad P(C|W) = \frac{P(C \cap W)}{P(W)} = \frac{\frac{3}{10}}{\frac{6}{10}} = 1/2$$

$$P(C|R \cap W) = \frac{P(C \cap R \cap W)}{P(R \cap W)} = \frac{\frac{1}{10}}{\frac{2}{10}} = 1/2$$

Addition Rule $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$$P(R \cup W) = P(R) + P(W) - P(R \cap W) = \frac{4}{10} + \frac{6}{10} - \frac{2}{10} = 8/10$$

$$P(C|R \cup W) = \frac{P(C(R \cup W))}{P(R \cup W)} = \frac{P(C|R)P(R) + P(C|W)P(W) - P(C|R \cap W)P(R \cap W)}{P(R \cup W)}$$

$$P(C|R \cup W) = \frac{(\frac{1}{2} \times \frac{4}{10}) + (\frac{1}{2} \times \frac{6}{10}) - (\frac{1}{2} \times \frac{2}{10})}{\frac{8}{10}} = \frac{\frac{2}{10} + \frac{3}{10} - \frac{1}{10}}{\frac{8}{10}} = \frac{\frac{4}{10}}{\frac{8}{10}} = \frac{1}{2}$$

Area Diagram

$P(R \cap C)$	$P(R \cap W \cap C)$	$P(W \cap C)$	$P(W \cap C)$	$P(C)$
$P(R)$	$P(R \cap W)$	$P(W)$	$P(W)$	$P(R^c \cap W^c \cap C^c)$

Section 2: Proof Practice

For this question please do the work (probably with pen and paper, although you should feel free to write a .Rmd solution if you're interested in the practice with), and then enter your solutions into the Gradescope assignment slot that is called "Unit 1 – Proof Practice."

Math proofs have a bad reputation: many of the instructors have negative memories of rote tasks about proving the Pythagorean Theorem, or some other such task. Perhaps you do too?

However, in this course and throughout many parts of data science the work that you will do will involve arguing for why the approach that you have taken is the *best* of the available options. A math proof is just one of these forms of argument.

For the question that follows, a compelling argument – a *proof* – should show the following two facts:

- That the value you arrive at is possible (for example, for the maximum possible value, you might state that the maximum possible value is 1); but, importantly,
- That there are no ways to arrive at a value that is higher (or lower as the case may be) than the value you arrive at.

Is part (b) that really makes this question a proof.

3. (6 points total) On the Overlap of Two Events

Suppose for events A and B , $P(A) = \frac{1}{2}$, $P(B) = \frac{3}{4}$, but we have no more information about the events.

- (3 points) What are the maximum and minimum possible values for $P(A \cap B)$?

Given that $P(A) = \frac{1}{2}$, $P(B) = \frac{3}{4}$ the maximum value of $P(A \cap B)$ is the smallest of the two probabilities.

Maximum value of $P(A \cap B) = \frac{1}{2}$

Given that any probability cannot exceed 1 and based on the maximum value of $P(A \cap B)$ the $P(A \cap B)^C = \frac{1}{2}$ we can assume that $P(A \cup B) = 1$.

Using the addition rule $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ we can find the minimum value of $P(A \cap B)$.

$$1 = \frac{1}{2} + \frac{3}{4} - P(A \cap B) = \frac{5}{4} - P(A \cap B)$$

$$P(A \cap B) = \frac{5}{4} - 1 = \frac{1}{4}$$

- (3 points) What are the maximum and minimum possible values for $P(A|B)$?

Given the limited information about $P(A)$ and the $P(B)$, we can find the maximum possible value for $P(A|B)$ using the conditional formula $P(A|B) = \frac{P(A \cap B)}{P(B)}$ and plug in the maximum possible value of the $P(A \cap B) = 1/2$ or the $P(A)$.

$$P(A|B) = \frac{P(A)}{P(B)} = \frac{\frac{1}{2}}{\frac{3}{4}} = 2/3$$

Using the conditional formula $P(A|B) = \frac{P(A \cap B)}{P(B)}$ and plugging in the minimum possible value of the $P(A \cap B) = 1/4$ we can determine the minimum possible value for $P(A|B)$ as:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{4}}{\frac{3}{4}} = 1/3$$