

Statistics for Data Science

Unit 3 Homework: Probability Theory

Brad Andersen
W203 Section 4

January 23, 2019

1. **Gas Station Analytics** At a certain gas station, 40% of customers use regular gas (event R), 35% use mid-grade (event M), and 25% use premium (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.

$$\begin{array}{lll} P(R) = 0.40 & P(M) = 0.35 & P(P) = 0.25 \\ P(F|R) = 0.30 & P(F|M) = 0.60 & P(F|P) = 0.50 \end{array}$$

Table 1: Known probabilities: Gas Station Analytics

- (a) What is the probability that the next customer will request regular gas and fill the tank?

$$P(F|R) = \frac{P(F \cap R)}{P(R)}, \text{ therefore } P(F \cap R) = P(R) \times P(F|R)$$

$$P(F \cap R) = 0.40 \times 0.30$$

$$P(F \cap R) = 0.12$$

Probability: 0.12

- (b) What is the probability that the next customer will fill the tank?

$$(P(F|R) \times P(R)) + (P(F|M) \times P(M)) + (P(F|P) \times P(P))$$

$$(0.30 \times 0.40) + (0.60 \times 0.35) + (0.50 \times 0.25)$$

$$(0.12) + (0.21) + (0.125)$$

Probability: 0.455

- (c) Given that the next customer fills the tank, what is the conditional probability that they use regular gas?

Let event T ("next customer fills tank") = 0.455.

$$P(R|T) = \frac{P(R \cap T)}{P(T)}, \text{ therefore } P(R|T) = \frac{P(T \cap R)}{P(T)}$$

$$P(R|T) = \frac{P(T|R) \times P(R)}{P(T)}$$

$$P(R|T) = \frac{0.30 \times 0.40}{0.455}$$

$$P(R|T) = \frac{0.12}{0.455}$$

Probability: 0.264

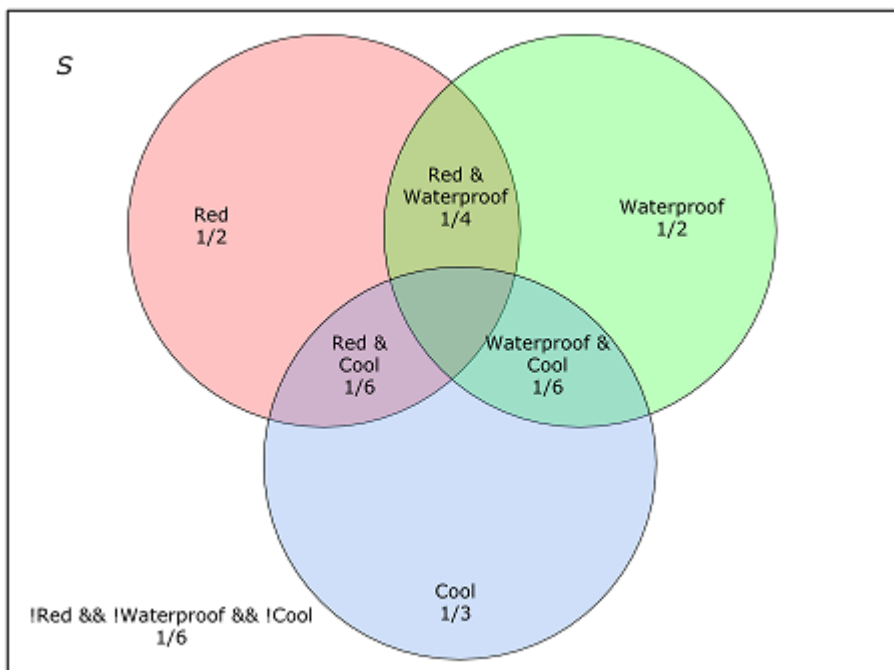
2. The Toy Bin

In a collection of toys, 1/2 are red, 1/2 are waterproof, and 1/3 are cool. 1/4 are red and waterproof. 1/6 are red and cool. 1/6 are waterproof and cool. 1/6 are neither red, waterproof, nor cool. Each toy has an equal chance of being selected.

$$P(R) = \text{red toy} \quad P(W) = \text{waterproof toy} \quad P(C) = \text{cool toy}$$

Table 2: Probability definitions: The Toy Bin

- (a) Draw an area diagram to represent these events.



- (b) What is the probability of getting a red, waterproof, cool toy?

Given that the probability of a toy being neither red nor waterproof nor cool is 1/6,

the probability for the union of the probabilities of a toy being red or waterproof or cool is $1 - 1/6$, or $5/6$. The value that must be calculated is the intersection of red, waterproof and cool in the Venn diagram, above.

$$\frac{5}{6} = P(R) + P(W) + P(C) - P(R \cap W) - P(R \cap C) - P(C \cap W) + P(R \cap W \cap C)$$

$$\frac{5}{6} = \frac{1}{2} + \frac{1}{2} + \frac{1}{3} - \frac{1}{4} - \frac{1}{6} - \frac{1}{6} + P(R \cap W \cap C)$$

$$\frac{5}{6} = \frac{3}{4} + P(R \cap W \cap C)$$

$$\frac{1}{12} = P(R \cap W \cap C)$$

Probability: $\underline{\underline{\frac{1}{12}}}$

- (c) 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?

$$\text{Probability} = P(R) - P(R \cap C)$$

$$\text{Probability} = \frac{1}{2} - \frac{1}{6}$$

$$\text{Probability} = \frac{6}{12} - \frac{2}{12}$$

$$\text{Probability} = \frac{4}{12} = \underline{\underline{\frac{1}{3}}}$$

- (d) Given that a randomly selected toy is red or waterproof, what is the probability that it is cool?

The following calculation assumes that "red or waterproof" does not mean that these are mutually exclusive events. Rather, I am interpreting "red or waterproof" to mean that the selected toy can be both red and waterproof.

$$\text{Probability} = P(R \cap C) + P(W \cap C) - P(R \cap W \cap C)$$

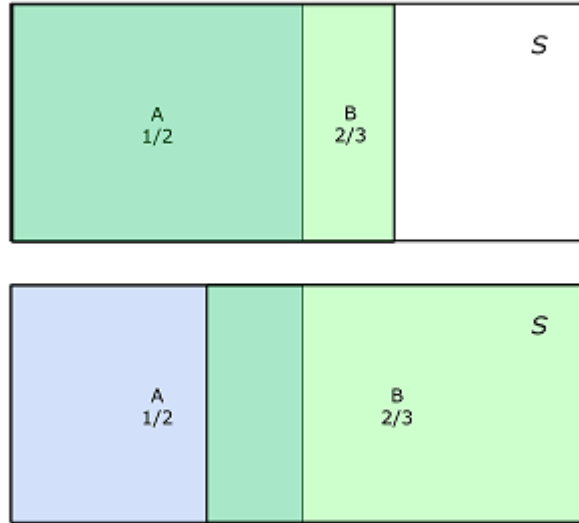
$$\text{Probability} = \frac{1}{6} + \frac{1}{6} - \frac{1}{12}$$

$$\text{Probability} = \frac{3}{12} = \underline{\underline{\frac{1}{4}}}$$

3. **On the Overlap of Two Events** Suppose for events A and B, $P(A) = 1/2$, $P(B) = 2/3$, but we have no more information about the events.

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

The following diagram represents the intersections of $P(A)$ and $P(B)$ as probabilities in S :



Given that the probability of S is 1, the maximum and minimum values for $P(A \cap B)$ are $\underline{\underline{\frac{1}{2}}}$ and $\underline{\underline{\frac{1}{6}}}$, respectively.

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

Given the definition of conditional probability:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Maximum and minimum values can be calculated as follows:

$$\text{Maximum: } P(A|B) = \frac{\frac{1}{2}}{\frac{2}{3}} = \underline{\underline{\frac{3}{4}}}$$

$$\text{Minimum: } P(A|B) = \frac{\frac{1}{6}}{\frac{2}{3}} = \underline{\underline{\frac{1}{4}}}$$

4. **Can't Please Everyone!** Among Berkeley students who have completed w203, 3/4 like statistics. Among Berkeley students who have not completed w203, only 1/4 like statistics. Assume that only 1 out of 100 Berkeley students completes w203. Given that a Berkeley student likes statistics, what is the probability that they have completed w203?

$$\begin{array}{lll} P(C) = \text{completed W203} & P(N) = \text{not completed W203} & P(L) = \text{like statistics} \\ P(C) = 0.01 & P(N) = 0.99 & \end{array}$$

Table 3: Probability definitions: Can't Please Everyone!

$$P(L|C) = \frac{P(L \cap C)}{P(C)}$$

$$P(L \cap C) = P(L|C) \times P(C)$$

$$P(L \cap C) = 0.75 \times 0.01$$

$$P(L \cap C) = 0.0075$$

$$\text{Probability} = \underline{\underline{0.0075}}$$