Bayes' Theorem

Example

- Three jars contain colored balls as described in the table below.
  - One jar is chosen at random and a ball is selected. If the ball is red, what is the probability that it came from the 2^{nd} jar?

Jar #	Red	White	Blue
1	3	4	1
2	1	2	3
3	4	3	2

Example

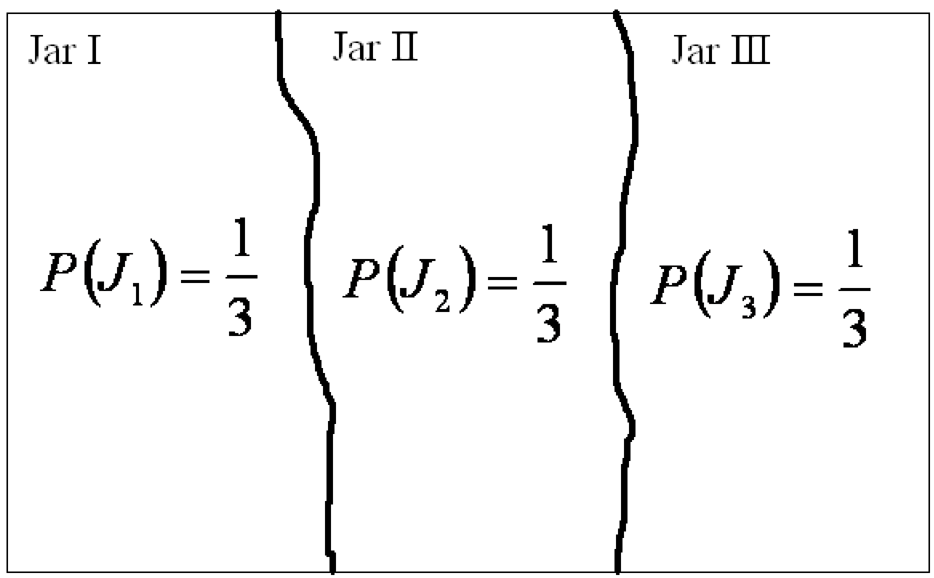
- We will define the following events:
  - J_1 is the event that *first* jar is chosen
  - J_2 is the event that *second* jar is chosen
  - J_3 is the event that *third* jar is chosen
  - R is the event that a *red* ball is selected

Example

- The events J_1, J_2, and J_3 mutually exclusive
  - Why?
    - *You can't chose two different jars at the same time*
- Because of this, our sample space has been divided or *partitioned* along these three events

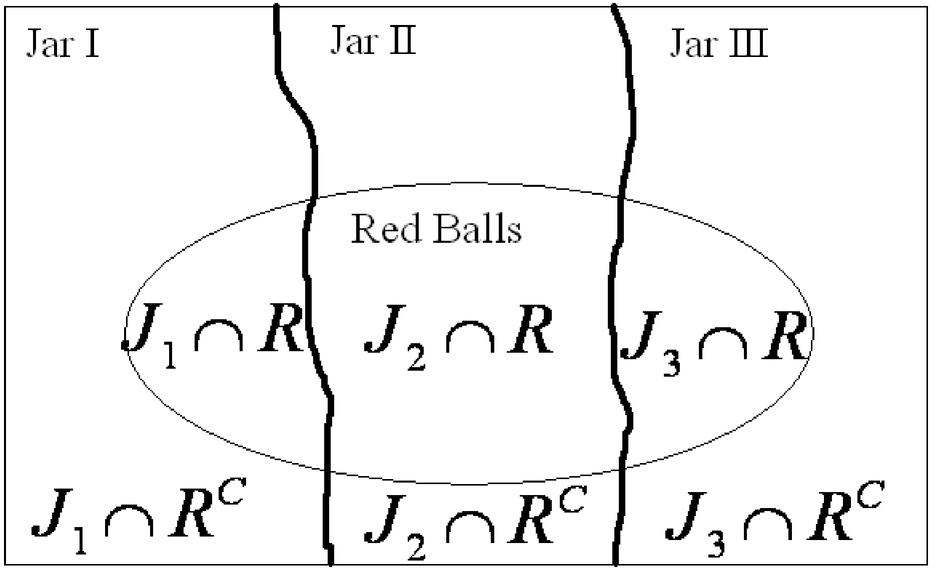
Venn Diagram

- Let’s look at the Venn Diagram



Venn Diagram

- All of the red balls are in the first, second, and third jar so their set overlaps all three sets of our partition



Finding Probabilities

- What are the probabilities for each of the events in our sample space?
- How do we find them?

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

Computing Probabilities

$$P(J_1 \cap R) = P(R \mid J_1)P(J_1) = \frac{3}{8} \cdot \frac{1}{3} = \frac{1}{8}$$

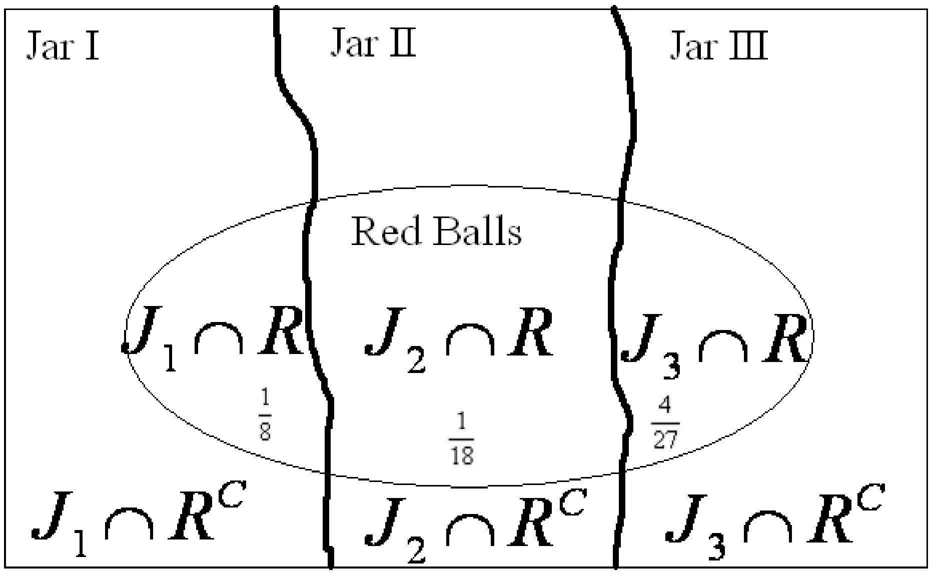
- Similar calculations show:

$$P(J_2 \cap R) = P(R \mid J_2)P(J_2) = \frac{1}{6} \cdot \frac{1}{3} = \frac{1}{18}$$

$$P(J_3 \cap R) = P(R \mid J_3)P(J_3) = \frac{4}{9} \cdot \frac{1}{3} = \frac{4}{27}$$

Venn Diagram

- Updating our Venn Diagram with these probabilities:



Where are we going with this?

- Our original problem was:
  - One jar is chosen at random and a ball is selected. If the ball is red, what is the probability that it came from the 2nd jar?
- In terms of the events we've defined we want:

$$P(J_2 \mid R) = \frac{P(J_2 \cap R)}{P(R)}$$

Finding our Probability

- We already know what the numerator portion is from our Venn Diagram
- What is the denominator portion?

$$P(J_2 \mid R) = \frac{P(J_2 \cap R)}{P(R)}$$

$$= \frac{P(J_2 \cap R)}{P(J_1 \cap R) + P(J_2 \cap R) + P(J_3 \cap R)}$$

Arithmetic!

- Plugging in the appropriate values:

$$P(J_2 \mid R) = \frac{P(J_2 \cap R)}{P(J_1 \cap R) + P(J_2 \cap R) + P(J_3 \cap R)}$$

$$= \frac{\left(\frac{1}{18}\right)}{\left(\frac{1}{8}\right) + \left(\frac{1}{18}\right) + \left(\frac{4}{27}\right)} = \frac{12}{71} \approx 0.17$$

Another Example—Tree Diagrams

All tractors made by a company are produced on one of three assembly lines, named Red, White, and Blue. The chances that a tractor will not start when it rolls off of a line are 6%, 11%, and 8% for lines Red, White, and Blue, respectively. 48% of the company's tractors are made on the Red line and 31% are made on the Blue line. What fraction of the company's tractors do not start when they roll off of an assembly line?

Define Events

- Let *R* be the event that the tractor was made by the red company
- Let *W* be the event that the tractor was made by the white company
- Let *B* be the event that the tractor was made by the blue company
- Let *D* be the event that the tractor won’t start

Extracting the Information

- In terms of probabilities for the events we've defined, this what we know:

$$P(R) = 0.48$$
$$P(W) = 0.21$$
$$P(B) = 0.31$$
$$P(D \mid R) = 0.06$$
$$P(D \mid W) = 0.11$$
$$P(D \mid B) = 0.08$$

What are we trying to find?

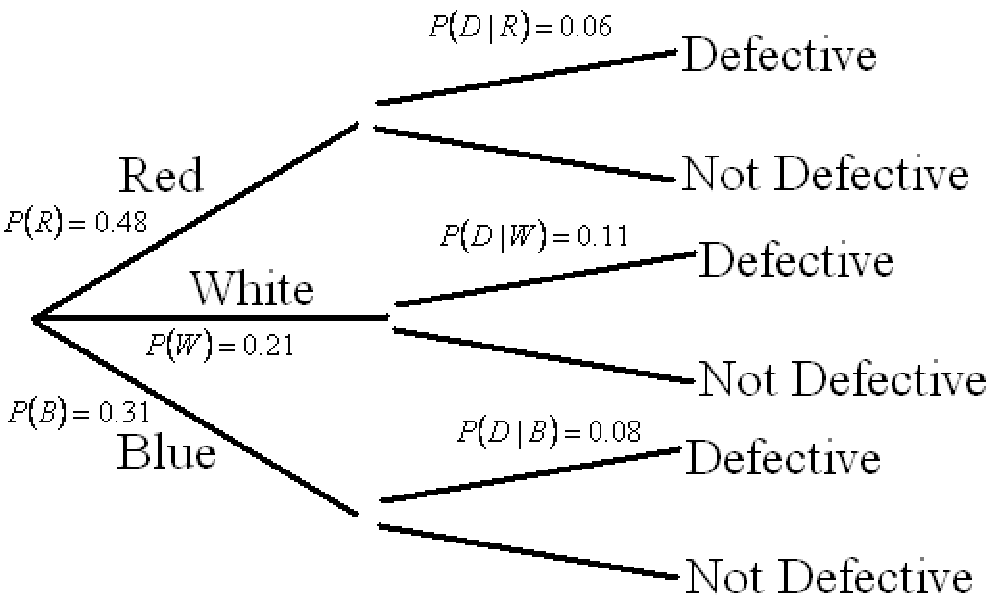
- Our problem asked for us to find:
  - The fraction of the company's tractors that do not start when rolled off the assembly line?
  - In other words:

$$P\left(D\right)$$

Tree Diagram

- Because there are three companies producing tractors we will divide or *partition* our sample space along those events only this time we'll be using a tree diagram

Tree Diagram



Follow the Branch?

- There are three ways for a tractor to be defective:
  - It was made by the Red Company
  - It was made by the White Company
  - It was made by the Blue Company

- To find all the defective ones, we need to know how many were:
  - Defective and made by the Red Company?
  - Defective and made by the White Company?
  - Defective and made by the Blue Company?

The Path Less Traveled?

- In terms of probabilities, we want:

$$P(R \cap D)$$

$$P(W \cap D)$$

$$P(B \cap D)$$

Computing Probabilities

- To find each of these probabilities we simply need to multiply the probabilities along each branch
- Doing this we find

$$P(R \cap D) = P(D \mid R)P(R)$$
$$P(W \cap D) = P(D \mid W)P(W)$$
$$P(B \cap D) = P(D \mid B)P(B)$$

Putting It All Together

- Because each of these events represents an instance where a tractor is defective to find the total probability that a tractor is defective, we simply add up all our probabilities:

$$P(D) = P(D \mid R)P(R) + P(D \mid W)P(W) + P(D \mid B)P(B)$$

Bonus Question:

- What is the probability that a tractor came from the red company given that it was defective?

$$P(R \mid D) = \frac{P(R \cap D)}{P(D)}$$

I thought this was called Bayes' Theorem?

- Bayes' Theorem
- Suppose that $B_{1,}$ $B_{2,}$ $B_{3,\ldots,}$ B_n partition the outcomes of an experiment and that A is another event. For any number, *k*, with $1 \le k \le n$, we have the formula:

$$P(B_k \mid A) = \frac{P(A \mid B_k) \cdot P(B_k)}{\sum_{i=1}^{n} P(A \mid B_i) \cdot P(B_i)}$$

In English Please?

- What does Bayes' Formula helps to find?
  - Helps us to find:

$$P\left(B \mid A\right)$$

  - By having already known:

$$P\left(A \mid B\right)$$