Mathematics for Analytics and Finance

Sami Najafi

MSIS2402/2405

Module 5

Events and their Probabilities

Events and Sample Space

Trial (Experiment): any procedure that can be infinitely repeated and has a well-defined set of possible outcomes known as the sample space.

Sample Space (Ω): the set of all possible outcomes of an experiment

- e.g., Experiment: Rolling a die.
  - Sample Space: $\Omega = \{1, 2, 3, 4, 5, 6\}$

- e.g., Experiment: Drawing a random card from a playing card deck
  - Sample Space: $\Omega = \{\text{All 52 cards of the deck}\}$

Event: is a subset of the sample space

- e.g., In rolling a die: $A = \{1, 2\}$ is the event that the outcome is 1 or 2

Events as Sets

Complement event (A^c): members of Ω that are not in A: $A^c = \{x \in \Omega \text{ and } x \notin A\}$

Impossible (null) event $(\emptyset)$: an empty set: $\emptyset = \{\}$

Union event $(A \cup B)$: any outcome is either in A or B (or both): $A \cup B = \{x \in A \text{ or } x \in B\}$

Joint event $(A \cap B)$: any outcome is in both A and B: $A \cap B = \{x \in A \text{ and } x \in B\}$

Disjoint events (mutually exclusive events): A and B have no intersections: $A \cap B = \emptyset$

Certain event: one of the outcomes in the event set will sure to occur: $A \cup B = \Omega$

Difference between events $(A - B)$: set of outcomes that are in A but not in B: $A - B = \{x \in A, x \notin B\}$

Sigma Algebra (σ-algebra)

σ-algebra: a collection F of subsets of Ω is called a σ-algebra if it satisfies the following:

- $\emptyset \in F$
- If $A_1, A_2, \ldots \in F$ then $\cup_{i=1}^{\infty} A_i \in F$
- If $A \in F$ then $A^c \in F$

Intuitively: F contains all possible information (i.e., events) that could materialize after conducting the experiment:

Example: Toss a coin once:

$$\Omega = \{H, T\}, \qquad F = \{\emptyset, \{H\}, \{T\}, \Omega\}$$

Example: Toss a coin twice:

$$\Omega = \{HH, HT, TH, TT\}$$

$$F = \{\emptyset, \{HH\}, \{HT\}, \{TH\}, \{TT\}, \{HH, HT\}, \{HH, TH\}, \ldots, \Omega\}$$

Probability Function

Definition: a probability function P on (Ω, F) is a function $P: F \to [0,1]$ satisfying:

- $P(\emptyset) = 0\,,\ P(\Omega) = 1$
- If $A_1, A_2, A_3, \dots$ are disjoints events in F, i.e., $A_i \cap A_j = \emptyset$ for all $i \neq j$, then:

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots$$

The triple (Ω, F, P) is called a probability space.

Remark: In many cases especially in games of chance it is natural to assume that all outcomes are equally likely. In such a case the probability of an event A is

$$\boxed{P(A) = \frac{n(A)}{n(\Omega)} = \frac{\text{Number of outcomes in } A}{\text{Number of outcomes in } \Omega} = \frac{\text{Number of desired outcomes}}{\text{Total number of outcomes}}}$$

Example

Find the probability of selecting a face card (Jack, Queen, or King) from a standard deck of 52 cards.

FC : The event that a face card is selected
Ω: The event that any card is selected

$$\mathrm{P}(A) = \frac{\mathrm{n}(FC)}{\mathrm{n}(\Omega)} = \frac{12}{52} = \frac{3}{13}$$

This is an example of a priori probability, which means it's calculated purely based on logical deduction and the known structure of the deck, without relying on past data. A priori probabilities are useful when you know the possible outcomes and their counts without needing historical data.

Example

Find the probability of selecting a male taking statistics from the population described in the below table:

	Taking Stats	Not Taking Stats	Total
Male	84	145	229
Female	76	134	210
Total	160	279	439

$$P(M) = \frac{n(M)}{n(\Omega)} = \frac{84}{439} \approx 0.191$$

This is an example of empirical probability, calculated from actual data rather than theoretical assumptions.

Empirical probability is often used in business when making predictions based on historical data.

Visualizing Events

Contingency Table: A contingency table (also known as a cross-tabulation or crosstab) is a matrix that displays the frequency distribution of two or more variables simultaneously. Each cell in the table shows the frequency (or the probability) of a particular combination of outcomes for the two variables (joint event)

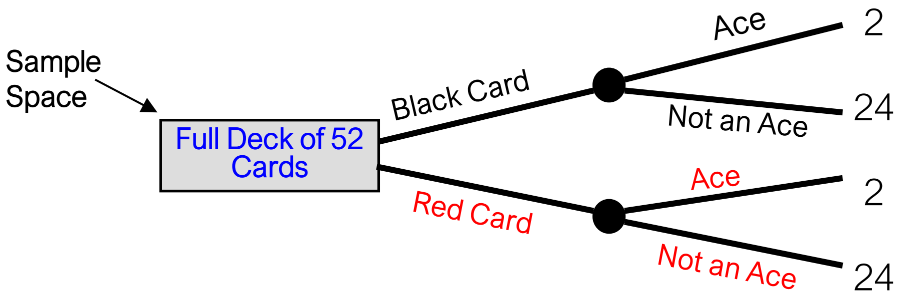
	Ace	Not Ace	Total
Black	2	24	26
Red	2	24	26
Total	4	48	52

Sample Space

Contingency tables are widely used in business analytics to examine relationships between two variables, like customer demographics (age, income) and purchase behavior (purchase, no purchase). They help identify patterns, such as whether a particular age group is more likely to buy a product.

Visualizing Events

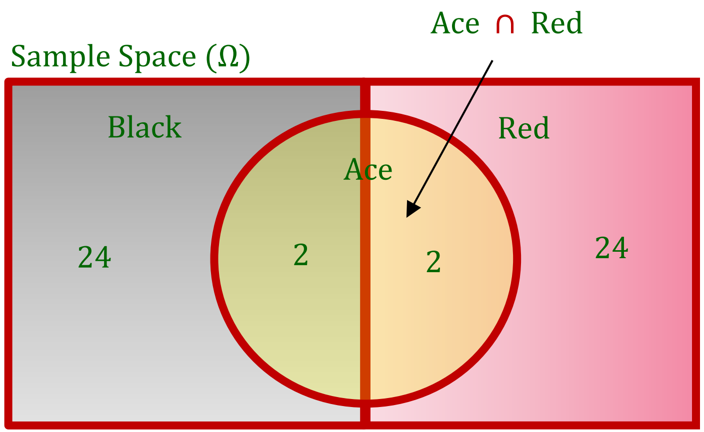
Decision Tree: A decision tree is a visual representation that breaks down events or decisions into a series of steps, using nodes and branches. It starts with a root node that represents the sample space. From the root node, branches extend to other nodes representing subsequent outcomes. The leaves or terminal nodes represent final outcomes, complete with associated frequencies or probabilities.



Decision trees are valuable for analyzing sequential decisions and conditional probabilities, common in business.

Visualizing Events

Venn Diagram: A Venn diagram is a visual tool to illustrate the relationships between different sets (or events) within a sample space. The sample space is shown with the outer rectangle and includes all the outcomes shown using common shapes (e.g., circles, rectangles, etc). Overlapping shapes are joint events.



Venn diagrams are often used in business to analyze overlapping customer segments, such as customers interested in both Product A and Product B.

Mutually Exclusive Events

Mutually exclusive (or disjoint) events: are events that cannot occur simultaneously.

Example: Drawing one card from a deck of cards

A = queen of diamonds; B = queen of clubs

- Events A and B are mutually exclusive: $P(A \cap B) = 0$

Collectively Exhaustive Events

Collectively exhaustive events: A set of events that together cover all possible outcomes of an experiment in the sample space. In other words, at least one event from the set must occur whenever the experiment is conducted.

Example:

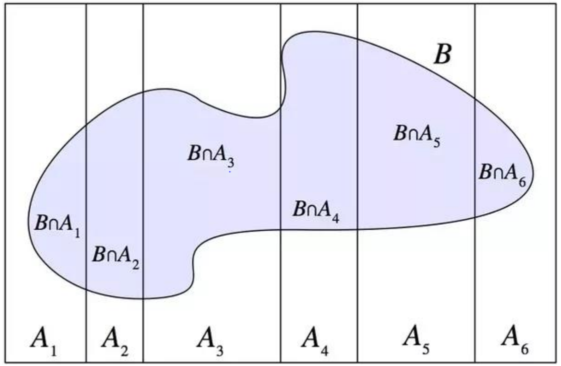
A = red cards; B = black cards;
C = diamonds; D = clubs

- Events A, B, C and D are collectively exhaustive (but not mutually exclusive – an ace may also be a diamond)
- Events A and B are collectively exhaustive and also mutually exclusive.
  - They partition the sample space into disjoint segments

Partitions of a Sample Space

Partitions (MECE events): Disjoint events whose union is the entire sample space

Example: A has partitioned the sample space in this figure



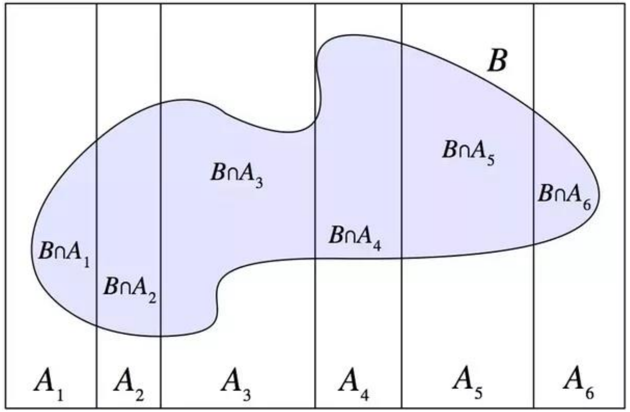
Complement as a Partition: The complement events A and A^c are partitions of a sample space and

$$P(A) = 1 - P(A^c)$$

Law of Total Probability

Law of Total Probability: If the events $A_1, A_2, \ldots, A_k$ partition the sample space, the probability of any event B can be found by summing the probabilities of B intersecting each partition A_i $(1 \leq i \leq k)$:

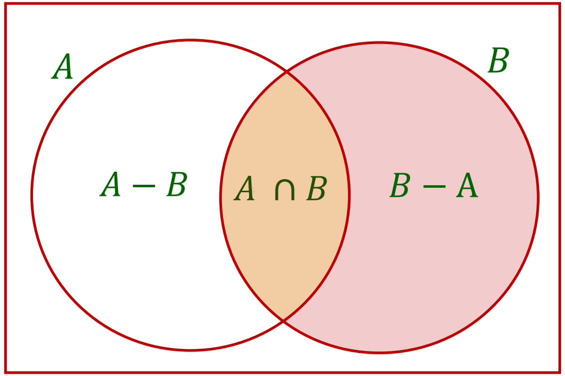
$$P(B) = \sum_{i=1}^{k} P(B \cap A_i) = P(B \cap A_1) + P(B \cap A_2) + \cdots + P(B \cap A_k)$$



General Addition Rule

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

Sample Space (Ω)



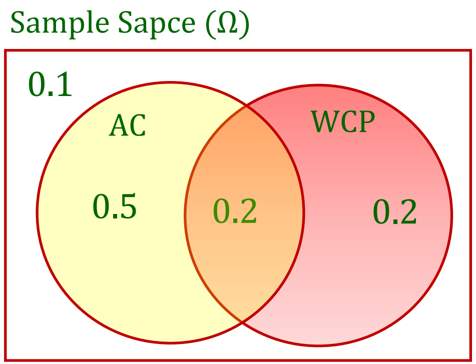
Because $A - B, A \cap B, B - A$ are disjoint events, by the definition of probability, we have:

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

Example

Of the cars on a used car lot, 70% have air conditioning (AC) and 40% have a wireless charging pad (WCP). 20% of the cars have both. What is the probability that a car has a CD player, given that it has WCP?

Solution



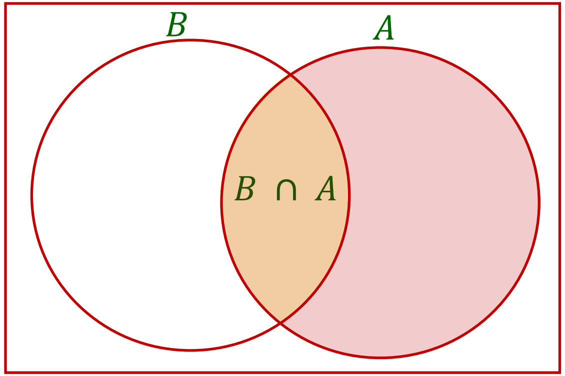
$$P(\mathrm{AC} \cup \mathrm{WCP}) = P(\mathrm{AC}) + P(\mathrm{WCP}) - P(\mathrm{AC} \cap \mathrm{WCP}) = 0.7 + 0.4 - 0.2 = 0.9$$

Conditional Probability

Conditional Probability

Definition: If $P(B) > 0$ then the conditional probability that A occurs given that B occurs is defined to be:

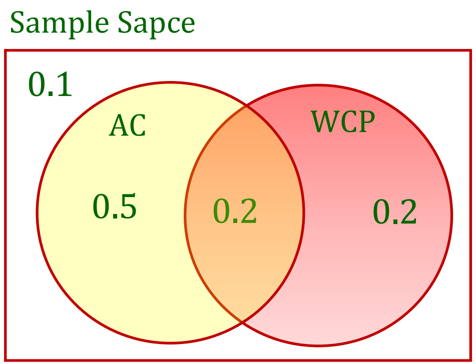
$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}}$$



Example

Of the cars on a used car lot, 70% have air conditioning (AC) and 40% have a wireless charging pad (WCP). 20% of the cars have both. What is the probability that a car has a CD player, given that it has WCP?

Solution



$$P(\text{WCP}|\text{AC}) = \frac{P(\text{WCP} \cap \text{AC})}{P(\text{AC})} = \frac{0.2}{0.7} = 0.2857$$

Example

A family has two children. What is the probability that both are boys, given that at least one is boy?

Solution

$$\Omega = \{\mathrm{BB}, \mathrm{BG}, \mathrm{GB}, \mathrm{GG}\}$$

$$P(\mathrm{BB}|\{\mathrm{BB}, \mathrm{BG}, \mathrm{GB}\}) = \frac{P(\mathrm{BB})}{P(\{\mathrm{BB}, \mathrm{BG}, \mathrm{GB}\})} = \frac{\mathrm{n}(\{\mathrm{BB}\})}{\mathrm{n}(\{\mathrm{BB}, \mathrm{BG}, \mathrm{GB}\})} = \frac{1}{3}$$

Independence

Definition: Events A and B are called independent if and only if the occurrence of one does not affect the probability of occurrence of the other.

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

Example: The two events Ace and Red are independent because:

$$P(\mathrm{Ace}|\mathrm{Red}) = P(\mathrm{Ace}) = \frac{2}{26}$$

Example: The two events AC and CD are dependent because:

$$P(\mathrm{CD}|\mathrm{AC}) = 0.28 \neq P(\mathrm{CD}) = 0.4$$

Remark: If A is independent of B, then B is also independent of A.

Multiplication Rule

(General) Multiplication Rule: For two events A and B, the probability that both events occur is given by:

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

Special Case for Independent Events: If A and B are independent then then the probability of both events happening simplifies to:

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

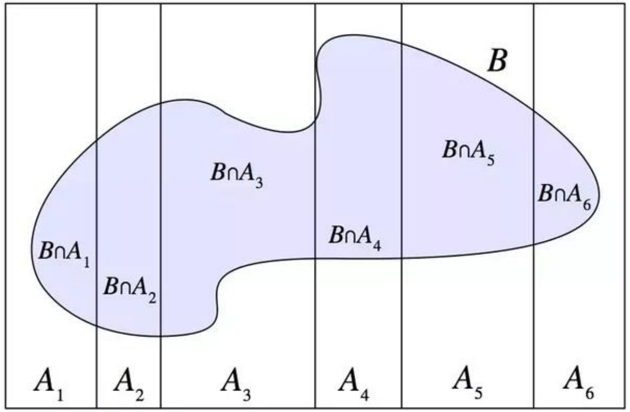
Extension to Multiple Independent Events: If we have multiple independent events $A_1, A_2, \ldots, A_k$ then the probability of all of them occurring is the product of their individual probabilities:

$$\mathrm{P}(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1) \times P(A_2) \times \cdots \times P(A_k)$$

Revisiting The Law of Total Probability

When $A_1, A_2, \dots, A_k$ are partitions of the sample space, the probability of any event B is the sum of the joint probabilities of B with each partition A_i $(1 \le i \le k)$:

$$P(B) = \sum_{i=1}^{k} P(B \cap A_i) = \sum_{i=1}^{k} P(B|A_i)P(A_i)$$



Example

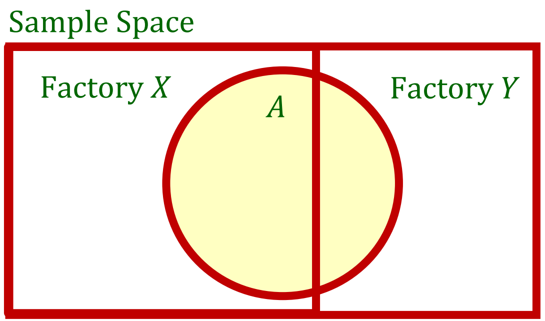
Suppose that there are two factories that supply light bulbs to the market. Factory X's bulbs have a track record of working for over 5000 hours in 99% of cases, while Factory Y's bulbs meet this criterion in 95% of cases. We also know that Factory X supplies 60% of the total bulbs available. What is the probability that a purchased bulb will work for longer than 5000 hours?

Solution

- A : the event that a bulb lasts over 5000 hrs
- X : the event that a bulb is from Factory X
- Y : the event that a bulb is from Factory Y

$$P(A) = P(X) \times P(A|X) + P(Y) \times P(A|Y)$$

$$= 0.6 \times 0.99 + 0.4 \times 0.95 = 0.974$$

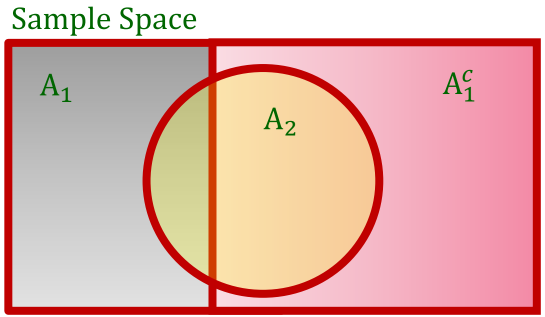


Example

When we draw two cards sequentially from a well shuffled deck, what is the probability of the second card being an ace?

Solution

- A_1 : the event that the first card is an ace
- A_2: the event that the second card is an ace



$$P(A_1) = P(A_1) \times P(A_2|A_1) + P(A_1^c) \times P(A_2|A_1^c)$$

$$= \frac{4}{52} \times \frac{3}{51} + \frac{48}{52} \times \frac{4}{51} = \frac{4}{52}$$

Example

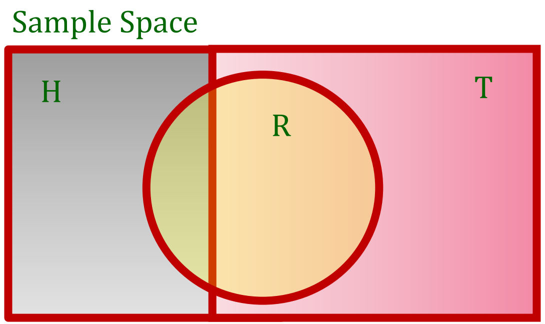
We have two hats: the first has 4 red balls and 6 green balls, and the second has 6 red and 4 green. We flip a fair coin: if it lands on heads, we randomly draw a ball from the first hat, but if it lands on tails, we draw from the second hat. What is the probability that we will draw a red ball?

Solution

- R : the event of drawing a red ball
- H : the event that the coin lands on heads (drawing from the first hat)
- T : the event that the coin lands on tails (drawing from the second hat)

$$P(R) = P(H) \times P(R|H) + P(T) \times P(R|T)$$

$$= \frac{1}{2} \times \frac{4}{10} + \frac{1}{2} \times \frac{6}{10} = \frac{1}{2}$$



Bayes' Theorem

Bayes' Formula

Bayes' Formula: allows us to calculate the probability of an underlying cause (event A_j) given that we observe a particular outcome (event B). This formula is widely used to update probabilities based on new information or evidence.

Formula: If $A_i, i = 1, \dots, k$ are partitions of sample space, each A_i having positive probability, then:

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^{k} P(B|A_i)P(A_i)}$$

Bayes' formula helps calculate the "posterior probability" of A_j after taking into account the evidence provided by B. This is essentially updating our prior belief $P(A_j)$ by weighting it with how likely B is under each scenario.

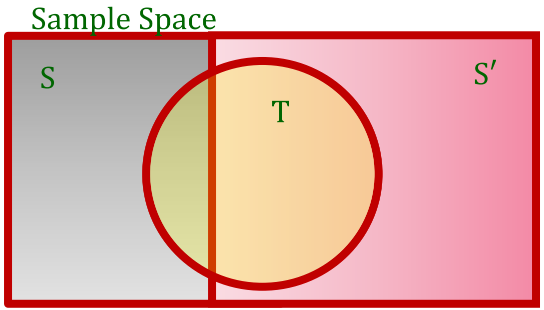
Example

A drilling company has provided an estimation of a 40% chance of striking oil for their newly drilled well. In order to gather more information, a diagnostic test has been scheduled. Based on historical data, it has been observed that 60% of successful wells have yielded positive test results, while 20% of unsuccessful wells have also yielded positive test results. Now, considering the fact that the test for this well has come back positive, what is the probability that the well will indeed strike oil and be classified as a successful well?

Solution

Consider the two events:

- $\mathrm{S} =$ the event that the well will strike oil (success)
- $T =$ the event that the test result is positive



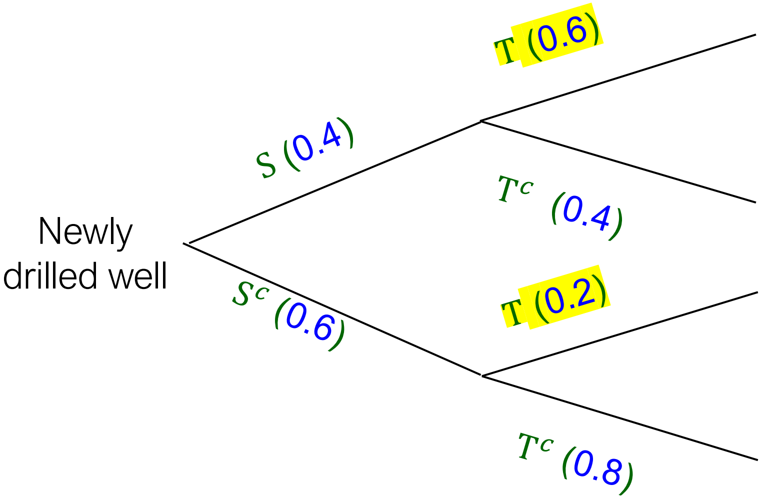
Then:

- $\mathrm{P(S) = 0.4,\ P(S^c) = 0.6}$
- $\mathrm{P(T|S) = 0.6,\ P(T|S^c) = 0.2}$

$$\mathrm{P(S|T) = \frac{P(S \cap T)}{P(T)} = \frac{P(S) \times P(T|S)}{P(S) \times P(T|S) + P(S^c) \times P(T|S^c)} = \frac{0.4 \times 0.6}{0.4 \times 0.6 + 0.6 \times 0.2} = 0.67}$$

- Therefore, receiving this good increased our belief about striking the oil successfully by about 27%.

Using Decision Trees



$$P(S \cap T) = 0.4 \times 0.6 = 0.24$$

$$P(S^c \cap T) = 0.6 \times 0.2 = 0.12$$

$$P(S|T) = \frac{P(S \cap T)}{P(T)} = \frac{P(S \cap T)}{P(S \cap T) + P(S^c \cap T)} = \frac{0.24}{0.24 + 0.12} = 0.67$$

Example

In previous cases, 40% of new-model televisions have been classified as successful, while 60% have been considered unsuccessful. Prior to introducing a new model television, the marketing research department conducts a comprehensive study and releases either a favorable or unfavorable report. In the past, 80% of the successful new-model televisions have received favorable market research reports, whereas 30% of the unsuccessful ones have received favorable reports. Now, for the new model of television currently being considered, the marketing research department has issued a favorable report. What is the probability that this television will be successful?

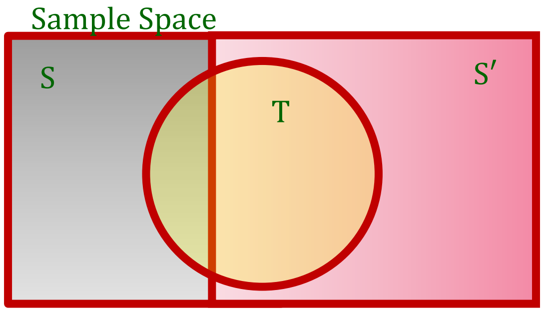
Solution

Consider the two events:

- $S = \{\text{successful TV model}\}$
- $T = \{\text{positive market report}\}$

Then:

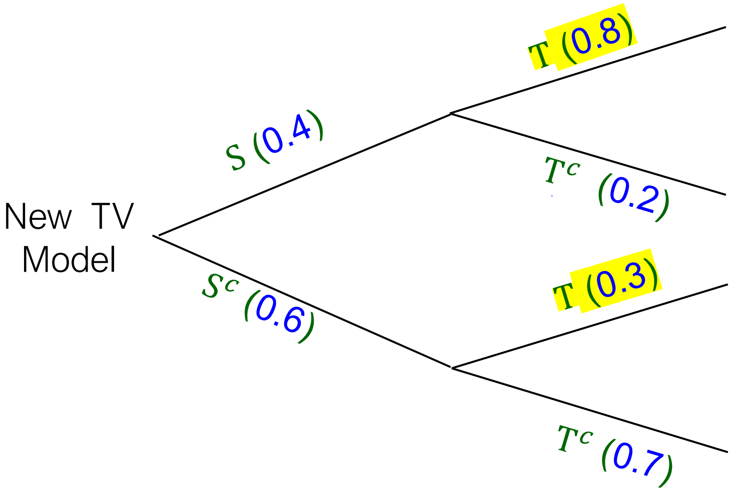
- $P(S) = 0.4,\ P(S') = 0.6$
- $P(T|S) = 0.8,\ P(T|S') = 0.3$



$$P(S|T) = \frac{P(S \cap T)}{P(T)} = \frac{P(S) \times P(T|S)}{P(S) \times P(T|S) + P(S') \times P(T|S')} = \frac{0.4 \times 0.8}{0.4 \times 0.8 + 0.6 \times 0.3} = 0.64$$

- The updated belief about the chance of the success of the new model after getting the new information is 64%.

Using Decision Trees



$P(S \cap T) = 0.4 \times 0.8 = 0.32$

$P(S^c \cap T) = 0.6 \times 0.3 = 0.18$

$$P(S|T) = \frac{P(S \cap T)}{P(T)} = \frac{P(S \cap T)}{P(S \cap T) + P(S^c \cap T)} = \frac{0.32}{0.32 + 0.18} = 0.64$$