Conditional Probability and Bayes Theorem

Probability Theory: Foundation for Data Science

with **Anne Dougherty**

Data Science

UNIVERSITY OF COLORADO **BOULDER**

Learning Goals for Module 2

In this module, we'll learn about conditional probability and Bayes formula. At the end of this module, learners should be able to:

- **Explain the concept of conditional probability.**
- **Calculate probabilities using conditioning and Bayes Theorem.**
- Explain the concepts of independence and mutually exclusive events and provide examples.
- See the relationship between conditional and independent events in a statistical experiment.

Conditional Probability

Suppose we have two events A and B from the same sample space S. We want to calculate the probability of event A, knowing that event B has occurred. B is the "conditioning event". Notation: $P(A|B)$, the probability of event A given that B has occurred.

Example: Roll a six-sided dice twice. Recall,
$S = \{(i,j) \mid i,j \in \{1,2,3,4,5,6\}\}$, $|S| = 36$ and each of the 36 outcomes of S is equally likely.

Let A be the event that at least one of the dice shows a 3.
$A = \{(3,1),(3,2),\ldots,(3,6),(1,3),(2,3),(4,3),(5,3),(6,3)\}$
$P(A) = 11/36$

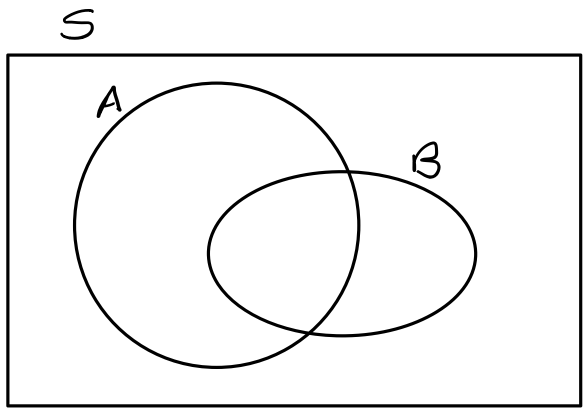
Let B be the event that the sum of the 2 dice is 9.
$B = \{(6,3),(3,6),(4,5),(5,4)\}$
$P(B) = 4/36$

Example continued

Question: Suppose we know that B has occurred. How does this change the probability of A? That is, find $P(A|B)$, the probability that at least one dice was a 3 given that the sum was 9.

- $P(A \cap B) =$

- $P(A|B) =$



Bayes Theorem

Conditional probability is defined as:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0$$

This leads to the **multiplication rule**

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

Bayes Theorem: Let $P(B) > 0$. Then,

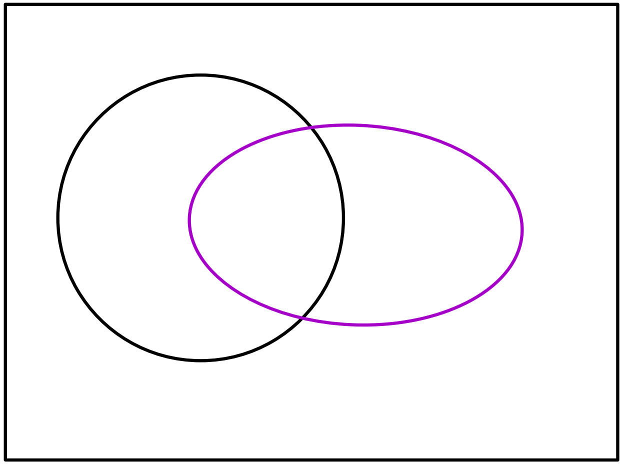
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Law of Total Probability

Given two events A and B from the same sample space,

$$B = (B \cap A) \cup (B \cap A^c)$$

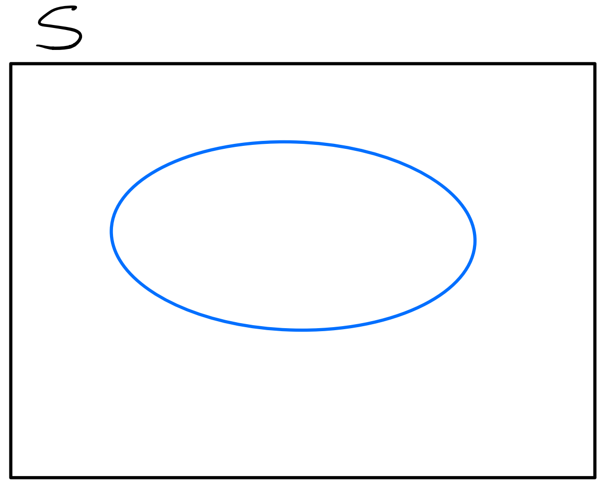
$$P(B) = P(B \cap A) + P(B \cap A^c) = P(B|A)P(A) + P(B|A^c)P(A^c)$$



Law of Total Probability - continued

Extend this idea to n sets $A_1, A_2, \ldots, A_n$ where $A_1 \cap \cdots \cap A_n = \emptyset$ and $\cup_{k=1}^{n} A_k = S$. Then,

$$P(B) = \sum_{k=1}^{n} P(B|A_k)P(A_k).$$



Example - Testing for a disease

Example: Suppose your company has developed a new test for a disease. Let event A be the event that a randomly selected individual has the disease and, from other data, you know that 1 in 1000 people has the disease. Thus, $P(A) = .001$. Let B be the event that a positive test result is received for the randomly selected individual. Your company collects data on their new test and finds the following:

- $P(B|A) = .99$
- $P(B^c|A) = .01$
- $P(B|A^c) = .02$

Calculate the probability that the person has the disease, given a positive test result. That is, find $P(A|B)$.

Example - continued

Example - Tree Diagram