Chapter 4

Conditional Probability & Independence

Warm Up

Suppose we toss a fair coin three times and I tell you that at least one of them landed heads. What is the probability that the other is a tail?

1. $S = \{HH, HT, TH, TT\}$
2. Three outcomes have at least one head, so the *reduced* sample space is $\{HH, HT, TH\}$.
3. Each outcome is still equally likely, and two of them have a tail.

$$P(\text{at least one tail}|\text{at least one head}) = 2/3$$

which is read as the probability of at least one tail *given* there is at least one head.

Warm Up

Suppose we toss a fair coin three times and I tell you that at least one of them landed heads. What is is the probability that the other is a tail?

We can also think of this in a different way. In our original sample space of four equally likely outcomes: $S = \{HH, HT, TH, TT\}$

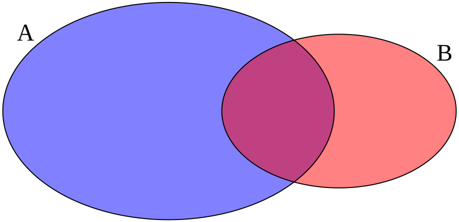
$$P(\text{at least 1 head}) = \frac{3}{4},$$
$$P(\text{ at least 1 tail and at least 1 head}) = \frac{2}{4}, \quad \textit{and} \quad \frac{2/4}{3/4} = \frac{2}{3};$$

so $2/3$ of the time when there is at least 1 head, there is also at least one tail.

$$P(\text{at least 1 tail}|\text{at least one head}) = 2/3$$

Visualizing

We are given that an element in B has occurred, and we wish to calculate the probability that it also belongs to the event A, that is, it belongs to $(A \cap B)$



The conditional probability is the ratio of $P(A \cap B)$ to $P(B)$.

Conditional probability definition

Definition 4.1 If A and B are events in S and $P(B) > 0$ then the **conditional probability** of A given B, written $P(A|B)$, is defined as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{1}$$

Example 4.1

The following table contains the prediction record of a TV weather forecaster for 100 days:

	Forecast		
Actual	Sunny	Cloudy	Total
Sunny	25	10	35
Cloudy	14	51	65
Total	39	61	100

Suppose we select a day at random[1]. Let A be the event that the forecast is for sunny weather and B the event that the actual weather is sunny. Determine the value of each of these probabilities.

a. $P(A)$

[1] this means each day is equally likely to be selected

Example 4.1

The following table contains the prediction record of a TV weather forecaster for 100 days:

	Forecast		
Actual	Sunny	Cloudy	Total
Sunny	25	10	35
Cloudy	14	51	65
Total	39	61	100

Suppose we select a day at random. Let A be the event that the forecast is for sunny weather and B the event that the actual weather is sunny. Determine the value of each of these probabilities.

b. $P(A^c)$

Example 4.1

The following table contains the prediction record of a TV weather forecaster for 100 days:

	Forecast		
Actual	Sunny	Cloudy	Total
Sunny	25	10	35
Cloudy	14	51	65
Total	39	61	100

Suppose we select a day at random. Let A be the event that the forecast is for sunny weather and B the event that the actual weather is sunny. Determine the value of each of these probabilities.

c. $P(B|A)$

Example 4.1

The following table contains the prediction record of a TV weather forecaster for 100 days:

	Forecast		
Actual	Sunny	Cloudy	Total
Sunny	25	10	35
Cloudy	14	51	65
Total	39	61	100

Suppose we select a day at random. Let A be the event that the forecast is for sunny weather and B the event that the actual weather is sunny. Determine the value of each of these probabilities.

d. $P(B^c|A)$

Example 4.1

The following table contains the prediction record of a TV weather forecaster for 100 days:

	Forecast		
Actual	Sunny	Cloudy	Total
Sunny	25	10	35
Cloudy	14	51	65
Total	39	61	100

Suppose we select a day at random. Let A be the event that the forecast is for sunny weather and B the event that the actual weather is sunny. Determine the value of each of these probabilities.

e. $P(B|A^c)$

Note that:

$$P(A^c|B) + P(A|B) = 1,$$

but

$$P(A|B) + P(A|B^c) \neq 1.$$

Chain rule for probabilities

Re-expressing the definition of a conditional probability

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

gives us a useful form for calculating intersection probabilities:

$$P(A \cap B) = P(A|B) \times P(B). \tag{4.2}$$

This is called the **chain rule for probabilities**. Likewise using

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

we can write

$$P(A \cap B) = P(B|A) \times P(A). \tag{4.3}$$

Equating equations (4.2) and (4.3) results in Bayes' theorem.

Bayes' theorem

Theorem 4.1 For events A and B:

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

Bayes' theorem is useful for calculating *inverse probabilities*.

Example 4.2

Suppose the probability of snow is 20% and that the probability of an accident on a snowy day is 40%, but only 2.5% on a non-snowy day. We select a day at random and learn there was an accident. What is the probability that there was snow involved?

Independent events

Definition 4.2 If

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

then we say that A and B are **independent events**.

- When A and B are independent events, the occurrence of one has no effect on the probability of the other.
    - In other words

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \times P(B)}{P(B)} = P(A). \tag{4.4}$$

- Similarly

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A) \times P(B)}{P(A)} = P(B). \tag{4.5}$$

- We can also simply use any of these equations as our definition of independence.

Example 4.3

A fair six sided dice is rolled twice. Suppose A is the event that the first throw yields a 2 or a 5 and B is the event that the sum of the two throws is 7.

a. Are A and B disjoint?

Example 4.3

A fair six sided dice is rolled twice. Suppose A is the event that the first throw yields a 2 or a 5 and B is the event that the sum of the two throws is 7.

b. Are A and B independent?

- If two events are disjoint, then they are definitely not independent!

Independence of more than two events

Definition 4.3 A collection of events $A_1, A_2, \ldots, A_n$ are said to be **mutually independent** if for *any* sub-collection $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$ we have:

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \times P(A_{i_2}) \times \cdots \times P(A_{i_k}).$$

- For example, three events A, B, C are mutually independent if:

$$\begin{aligned}
P(A \cap B) &= P(A) \cdot P(B), \\
P(B \cap C) &= P(B) \cdot P(C), \\
P(A \cap C) &= P(A) \cdot P(C), \\
P(A \cap B \cap C) &= P(A) \cdot P(B) \cdot P(C).
\end{aligned}$$

Example 4.4

A system composed of n separate components is said to be a parallel system if it functions when at least one of the components function. For such a system, if component i, which is independent of the other components, functions with probability $p_i, i = 1, \ldots, n$, what is the probability that the system fails to function?