Chapter 4

Conditional Probability & Independence

Warm Up

Suppose we toss a fair coin ~~three~~ two 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~~ two 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: $S = \{HH, HT, TH, TT\}$

$$P(\text{at least 1 head}) = \frac{3}{4},$$

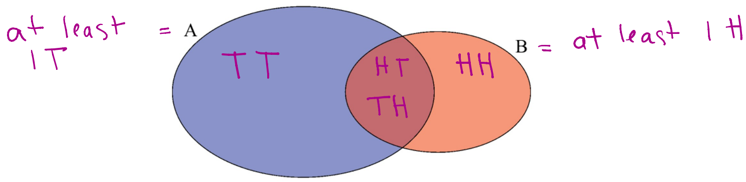
$$P(\text{ at least 1 tail and at least 1 head}) = \frac{2}{4}, \quad \textit{and} \quad \frac{2/4}{3/4} = \frac{2}{3};$$

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

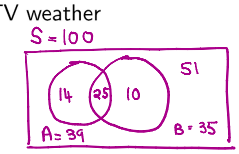
$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \quad (1)$$

Example 4.1

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

A A^c

	Forecast		
Actual	Sunny	Cloudy	Total
B → Sunny	25	10	35
B^c → 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. $P(A) = \frac{39}{100} = .39$

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

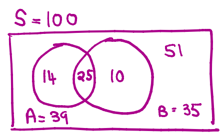
Example 4.1

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. $P(A^c) = 1 - P(A) = 1 - .39 = 0.61$



Example 4.1

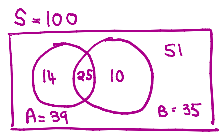
The following table contains the prediction record of a TV weather forecaster for 100 days:

	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. $P(B|A) = \frac{25}{39}$

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{25/100}{39/100} = \frac{25}{39}$$



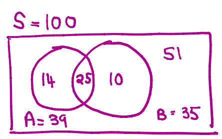
Example 4.1

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

	Actual	Forecast A: Sunny	A^c: Cloudy	Total
B	Sunny	25	10	35
B^c	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.

a. $P(B^c|A) = \frac{14}{39}$



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

$\frac{25}{39} + \frac{14}{39}$

Example 4.1

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

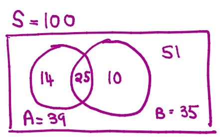
		A	A^c	
		Forecast		
	Actual	Sunny	Cloudy	Total
B	Sunny	25	10	35
B^c	Cloudy	14	51	65
	Total	39	61	100

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. $P(B|A^c) = \frac{10}{61}$

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

$\frac{10}{61} + \frac{14}{39} \neq 1$



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*.

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

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?

Event: A = snow on a randomly selected day

Given $P(A) = .2$

B = accident on a randomly " "

Given: $P(B|A) = .4$

$P(B|A^c) = .025$

Want $P(A|B)$

Prob. tree for bookkeeping

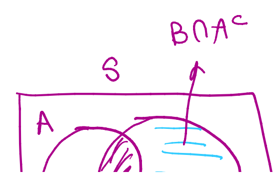
(A) Snow?

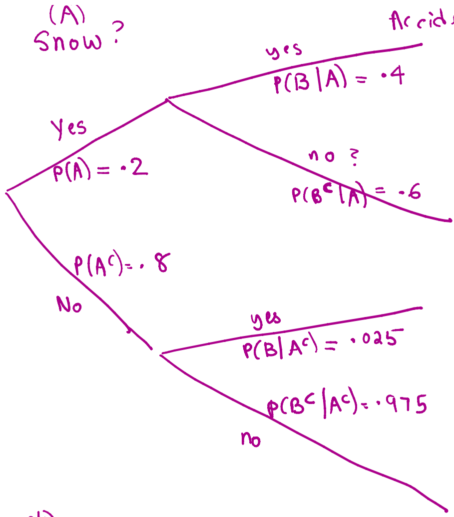
(B) Accident?

yes

$P(B|A) = .4$

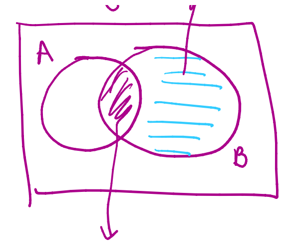
$P(B|A) = \frac{P(B \cap A)}{P(A)} \Rightarrow P(B \cap A) = .4 \times .2 = .08$ ✓





$P(B|A) = \frac{P(B \cap A)}{P(A)} \Rightarrow P(B \cap A) = .4 \times .2 = .08$ ✓

$P(B|A^c) = \frac{P(B \cap A^c)}{P(A^c)} \Rightarrow P(B \cap A^c) = .025 \times .8$ ✓
$= 0.02$



$P(A \cap B) = .08$

$P(B \cap A^c) = 0.02$

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

$P(B) = .08 + .02$
$= \boxed{0.1}$

want)

$$P(A|B) = \frac{P(B|A).P(A)}{P(B)}$$ Bayes' Rule

$$= \frac{.4 \times .2}{P(B)}$$

$$= \frac{.4 \times .2}{0.1}$$

Independent events

$P(A \cap B) = P(A|B).P(B)$
$= P(B|A).P(A)$ } chain rule

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?

$S = \{(1,1), (1,2), \dots, (1,6), (2,1), (2,2), \dots, (2,6), \dots, (6,1), (6,2), \dots, (6,6)\}$

Mult. rule

$|S| = 36$ (equally likely outcomes) $\frac{6}{1} \cdot \frac{6}{2} = 36$

$A = \{(2,1), (2,2), \dots, (2,6), (5,1), (5,2) \dots, (5,6)\}$

$P(A) = \frac{|A|}{|S|}$ (equally likely rule)

$= 12/36 = \boxed{1/3}$

$B = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$

$P(B) = \frac{|B|}{} = \frac{6}{} = \boxed{\frac{1}{6}}$

$B = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$

$P(B) = \frac{|B|}{|S|} = \frac{6}{36} = \boxed{\frac{1}{6}}$

Are A & B disjoint? Is $A \cap B = \emptyset$?

$A \cap B = \{(2,5), (5,2)\} \neq \emptyset$

$\therefore$ A and B are not 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?

A and B are independent if $P(A \cap B) = P(A) \cdot P(B)$

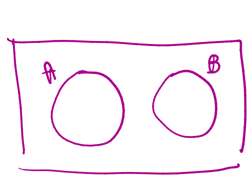
We know $P(A) = \frac{1}{3}$, $P(B) = \frac{1}{6}$, $P(A \cap B) = \frac{2}{36} = \frac{1}{18}$

Since $\frac{1}{18} = \frac{1}{3} \times \frac{1}{6}$

yes A and B are independent.

(Independence is hard to visualize with a Venn diagram)

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



$A \cap B = \emptyset$

Pf $P(A \cap B) = P(\emptyset) = 0$

But $P(A) \neq 0$ and $P(B) \neq 0$

$\therefore P(A \cap B) \neq P(A) \cdot P(B)$

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:

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

For 4 events A, B, C, D need

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

$P(A \cap C) = P(A) \cdot P(C)$ $+$ $P(A \cap B \cap C) = P(A)\, P(B)\, P(C)$

$P(A \cap B \cap D) = P(A)\, P(B)\, P(D)$ $+$ $P(A \cap B \cap C \cap D) = P(A) P(B) P(C) P(D)$

$P(A \cap B) = P(A) \cdot P(B)$
$P(A \cap C) = P(A) \cdot P(C)$
$P(A \cap D) = P(A) \cdot P(D)$
$P(B \cap C) = P(B) \cdot P(C)$
$P(B \cap D) = P(B) \cdot P(D)$
$P(C \cap D) = P(C) \cdot P(D)$

$+$

$P(A \cap B \cap C) = P(A)\, P(B)\, P(C)$
$P(A \cap B \cap D) = P(A)\, P(B)\, P(D)$
$P(A \cap C \cap D) = P(A)\, P(C)\, P(D)$
$P(B \cap C \cap D) = P(B)\, P(C)\, P(D)$

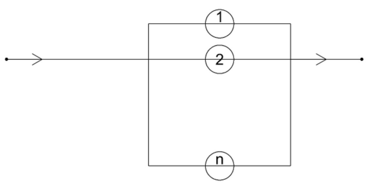
$+$

$P(A \cap B \cap C \cap D) = P(A) P(B) P(C) P(D)$

for A, B, C, D to be independent.

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?



A_i: relay i fails to function

$P(A_i) = 1 - P(A_i^c) = 1 - p_i$

$\hookrightarrow$ relay i functions

$P(\text{system fails}) = P(\text{all } n \text{ relays fail})$

$= P(A_1 \cap A_2 \cap A_3 \ldots \cap A_n)$

$= P(A_1) \times P(A_2) \times \ldots \times P(A_n)$ [since each relay operates independently]

$= (1 - p_1) \times (1 - p_2) \times \cdots \times (1 - p_n)$