Probability and Random Variables (ECE313/ECE317)

Fatima Taousser

Departement of Electrical Engineering and Computer Sciences, UTK
ftaousse@utk.edu

Conditional Probability
Fall 2023

Probability: Conditioning

- Conditional probability: The conditional probability is defined as

$$\mathbb{P}(A|B) = \text{ probability of } A, \text{ given that event } B \textbf{ occurred or certain}.$$

$$\Rightarrow \text{ use } \textbf{new information} \text{ to revise a model}$$

- B becomes our new universe (we are **certain** that B occurs)

Example: Consider the weather of the 6th of February of the last 10 years

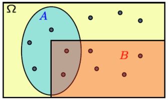
Year	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	
Raining	X	X		X	X	X	X		X	X	8 days
Windy		X	X	X		X		X	X	X	7 days
Humid	X	X		X			X	X			5 days

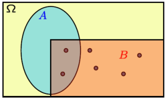
$$\mathbb{P}[Raining] = \frac{8}{10} = 80\%, \ \mathbb{P}[Raining|Windy] = \frac{5}{7} = 71\%, \ \mathbb{P}[Raining|windy \cap Humid] = \frac{2}{3} = 66\%$$

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

- Assumption: $\mathbb{P}(B) \neq 0$

Probability: Conditioning

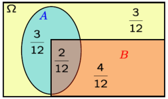


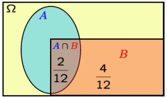


$$|\Omega| = 12, \quad |A| = 5, \quad |B| = 6, \quad |A \cap B| = 2$$

$$\mathbb{P}(A) = \frac{5}{12}, \quad \mathbb{P}(B) = \frac{6}{12}, \quad \mathbb{P}(A \cap B) = \frac{2}{12}$$

- Consider that B is the **new universe** $\Rightarrow \mathbb{P}(A|B) = \frac{2}{6}$ and $\mathbb{P}(B|B) = 1$





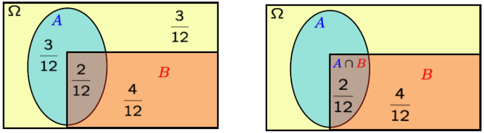
$$\mathbb{P}(A|B) = \frac{\frac{2}{12}}{\frac{6}{12}} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{2}{12} \times \frac{12}{6} = \frac{2}{6}$$

Probability: Conditioning

- Consequence (symmetry):

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

- Consider the previous example:



$$\mathbb{P}(B|A) = \frac{\frac{2}{12}}{\frac{5}{12}} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{2}{5}$$

Probability: Conditioning

Example: Consider the two rolls of a tetrahedral die

$\Omega = \{(1,1),(1,2),\ldots(4,4)\} = 16$ elements

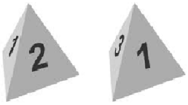
- Event $B = \{\min(X,Y) = 2\}$

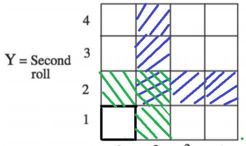
$$\Rightarrow B = \{(2,2),(2,3),(2,4),(3,2),(4,2)\} \;\rightarrow\; |B| = 5 \Rightarrow \mathbb{P}(B) = \frac{5}{16}$$

- Event: $M = \{\max(X,Y) = 2\}$

$$\Rightarrow M = \{(1,2),(2,1),(2,2)\} \;\rightarrow\; |M| = 3 \Rightarrow \mathbb{P}(M) = \frac{3}{16}$$

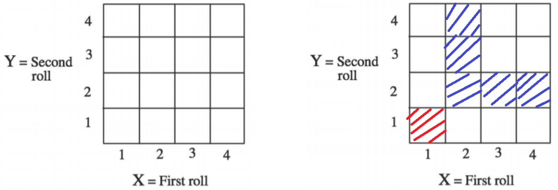
$$\mathbb{P}(M|B) = \frac{\mathbb{P}(M \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(2,2)}{\mathbb{P}(B)} = \frac{\frac{1}{16}}{\frac{5}{16}} = \frac{1}{5}$$





Probability: Conditioning

Example



- Event $B = \{\min(X, Y) = 2\}$
$= \{(2,2),(2,3),(2,4),(3,2),(4,2)\} \;\rightarrow\; |B| = 5 \Rightarrow \mathbb{P}(B) = \frac{5}{16}$
- Event $M = \{\max(X, Y) = 1\} = \{(1,1)\} \;\rightarrow\; |M| = 1 \Rightarrow \mathbb{P}(M) = \frac{1}{16}$

$\rightarrow \;\; M \cap B = \emptyset$

$$\mathbb{P}(M|B) = \frac{\mathbb{P}(M \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\emptyset)}{\mathbb{P}(B)} = 0$$

Probability: Model based on conditional probability

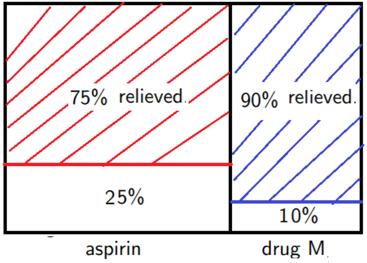
Example:

For headaches three out of five patients take aspirin (or equivalent), two out of five take a drug M.

-With aspirin, 75% of the patients have been relieved.

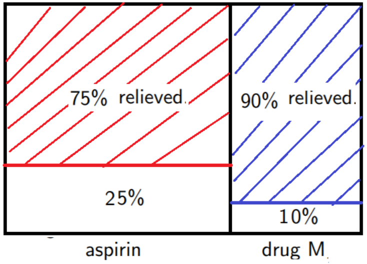
-With drug M, 90% of the patients have been relieved.

1. What is the overall rate of people relieved?
2. What is the likelihood that a patient has taken aspirin knowing that he has been relieved?



Probability: Model based on conditional probability

Modeling the problem:



-Universe: $\Omega = \{5 \text{ person have headaches}\}$

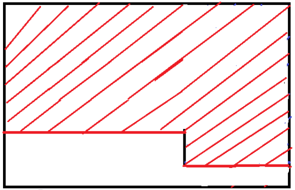
- Event A: $\{3 \text{ patients took aspirin}\} \rightarrow \mathbb{P}(A) = \frac{3}{5}$
- Event B: $\{2 \text{ patients took drug M}\} \rightarrow \mathbb{P}(B) = \frac{2}{5}$
- Event C: $\{\text{Patient is relieved}\}$
- We have $\mathbb{P}(C|A) = 0.75$
- We have $\mathbb{P}(C|B) = 0.9$

Probability: Model based on conditional probability

Solution: 1) The overall rate of people who relieved $= \mathbb{P}(C)$.

- In this example, we have $A \cup B = \Omega$ and $A \cap B = \emptyset$
- The event C can be written as: $C = \Omega \cap C = (A \cup B) \cap C = (A \cap C) \cup (B \cap C)$

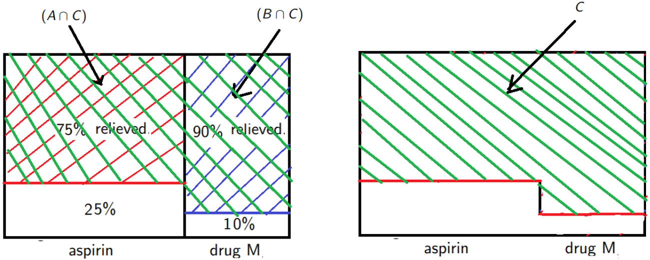
$$\begin{aligned}
\mathbb{P}(C) &= \mathbb{P}[(A \cap C) \cup (B \cap C)] \\
&= \mathbb{P}(A \cap C) + \mathbb{P}(B \cap C) - \mathbb{P}(A \cap B \cap C) \\
&= \mathbb{P}(A \cap C) + \mathbb{P}(B \cap C) - \mathbb{P}(\emptyset \cap C) \\
&= \mathbb{P}(A \cap C) + \mathbb{P}(B \cap C) - \mathbb{P}(\emptyset) \\
&= \mathbb{P}(A \cap C) + \mathbb{P}(B \cap C) \\
&= \mathbb{P}(A)\mathbb{P}(C|A) + \mathbb{P}(B)\mathbb{P}(C|B) \\
&= \tfrac{3}{5}(0.75) + \tfrac{2}{5}(0.9) = 0.81.
\end{aligned}$$



Probability: Model based on conditional probability

2) The likelihood that a patient has taken aspirin knowing that he has been relieved?
$\rightarrow \ \mathbb{P}(A|C)$

$$\mathbb{P}(A|C) = \frac{\mathbb{P}(A \cap C)}{\mathbb{P}(C)} = \frac{\mathbb{P}(A)\ \mathbb{P}(C|A)}{\mathbb{P}(C)} = \frac{\frac{3}{5}(0.75)}{0.81} = 0.5556$$



2) The likelihood that a patient has taken drug M knowing that he has been relieved

$$\rightarrow \ \mathbb{P}(B|C) = \frac{\mathbb{P}(B \cap C)}{\mathbb{P}(C)} = \frac{\mathbb{P}(B)\ \mathbb{P}(C|B)}{\mathbb{P}(C)} = \frac{\frac{2}{5}(0.9)}{0.81} = 0.4444$$

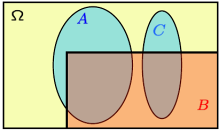
Probability: Conditioning

- Properties of conditional probability:

- $\mathbb{P}(A|\Omega) = \dfrac{\mathbb{P}(A \cap \Omega)}{\mathbb{P}(\Omega)} = \dfrac{\mathbb{P}(A)}{\mathbb{P}(\Omega)} = \dfrac{\mathbb{P}(A)}{1} = \mathbb{P}(A)$ and $\mathbb{P}(B|\Omega) = \mathbb{P}(B)$
- $\mathbb{P}(\Omega|B) = \dfrac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \dfrac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$
- $\mathbb{P}(B|B) = \dfrac{\mathbb{P}(B \cap B)}{\mathbb{P}(B)} = \dfrac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$
- $\mathbb{P}(A \cap B) = \mathbb{P}(B)\mathbb{P}(A|B)$ and $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B|A)$

Probability: Conditioning

- If $A \cap C = \emptyset$, we have $\mathbb{P}((A \cup C)|B) = \mathbb{P}(A|B) + \mathbb{P}(C|B)$



$$
\begin{aligned}
\mathbb{P}(A \cup C|B) &= \frac{\mathbb{P}(A \cup C) \cap B}{\mathbb{P}(B)} \\
&= \frac{\mathbb{P}[(A \cap B) \cup (C \cap B)]}{\mathbb{P}(B)} \\
&= \frac{\mathbb{P}(A \cap B) + \mathbb{P}(C \cap B) - \mathbb{P}[(A \cap B) \cap (C \cap B)]}{\mathbb{P}(B)} \\
&= \frac{\mathbb{P}(A \cap B) + \mathbb{P}(C \cap B) - \mathbb{P}(\emptyset)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} + \frac{\mathbb{P}(C \cap B)}{\mathbb{P}(B)} \\
&= \mathbb{P}(A|B) + \mathbb{P}(C|B)
\end{aligned}
$$

Consequence: $A \cap A^c = \emptyset \Rightarrow \mathbb{P}(\underbrace{(A \cup A^c)}_{\Omega}|B) = \mathbb{P}(A|B) + \mathbb{P}(A^c|B) = 1$

$\Rightarrow \mathbb{P}(A|B) = 1 - \mathbb{P}(A^c|B)$

Probability: Radar model based on conditional probability

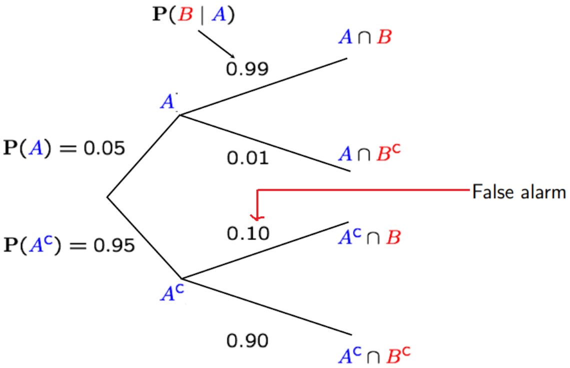
- Event A: An airplane is flying above
  $\rightarrow$ A^c: Nothing is flying above.
- Event B: Something registers on the radar's screens
  $\rightarrow$ B^c : The radar is not detecting anything.
- Let $\mathbb{P}(A) = 0.05$
- Let $\mathbb{P}(B|A) = 0.99$ $\quad \Rightarrow \mathbb{P}(B^c|A) = 1 - 0.99 = 0.01$
- Let $\mathbb{P}(B|A^c) = 0.1$ $\rightarrow$ False alarm $\quad \Rightarrow \mathbb{P}(B^c|A^c) = 1 - 0.1 = 0.90$

- Question:
What is the probability that an airplane is flying above when something registers on the radar's screen (we want to check the reliability of the radar).

$$\mathbb{P}(A|B) = ?$$

Probability: Model based on conditional probability

- Event A: An airplane is flying above
- Event B: Something registers on the radar screens



Probability: Model based on conditional probability

- By giving a conditional probability, can we compute $\mathbb{P}(A \cap B)$ and $\mathbb{P}(B)$?

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \Rightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A).\mathbb{P}(B|A) = (0.05)(0.99) = 0.0495$$

$$\mathbb{P}(A^c \cap B) = \mathbb{P}(A^c).\mathbb{P}(B|A^c) = (0.95)(0.1) = 0.095$$

$$\mathbb{P}(B) = \mathbb{P}((A \cup A^c) \cap B) = \mathbb{P}((A \cap B) \cup (A^c \cap B)) = \mathbb{P}(A \cap B) + \mathbb{P}(A^c \cap B)$$

$$= 0.0495 + 0.095 = 0.1445$$

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{0.0495}{0.1445} = \underline{0.342}$$

- The radar is **not reliable**: "Most of the time there is nothing but the radar detect a flying plane with a rate of 10% " $\rightarrow$ "false alarms are pretty common"

Probability: Bayes's rule:(Thomas Bayes, British Mathematician, 1701-1761)

Baye's rule:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A).\mathbb{P}(A)}{\mathbb{P}(B)} \quad \text{and} \quad \mathbb{P}(B|A) = \frac{\mathbb{P}(A|B).\mathbb{P}(B)}{\mathbb{P}(A)}$$

Proof:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \Rightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A|B).\mathbb{P}(B) \qquad (1)$$

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \Rightarrow \mathbb{P}(A \cap B) = \mathbb{P}(B|A).\mathbb{P}(A) \qquad (2)$$

From (1) and (2), we get

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A).\mathbb{P}(A)}{\mathbb{P}(B)} \quad \text{and} \quad \mathbb{P}(B|A) = \frac{\mathbb{P}(A|B).\mathbb{P}(B)}{\mathbb{P}(A)}$$

Probability: Bayes's rule

- Provide us a way to update our beliefs based on the arrival of new, relevant pieces of evidence

$\Rightarrow$ use prior knowledge to improve our probability estimation.

Example

- Application to the plane and radar example:

- we know $\mathbb{P}(A)$ (prior probabilities)
- we know $\mathbb{P}(B|A) \rightarrow$ new information
- we know $\mathbb{P}(B|A^c) \rightarrow$ new information
- we computed the $\mathbb{P}(B)$
- we want to compute $\mathbb{P}(A|B)$

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A).\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{(0.99) \times (0.05)}{0.1445} = 0.342.$$

Probability: Bayes's rule

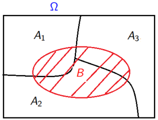
Example

- Application to the headache's drugs example:
  - We know $\mathbb{P}(A)$ (prior probabilities)
  - We know $\mathbb{P}(B)$ (prior probabilities)
  - We know $\mathbb{P}(C|A) \rightarrow$ new information
  - We know $\mathbb{P}(C|B) \rightarrow$ new information
  - We computed the $\mathbb{P}(C)$
  - We want to compute $\mathbb{P}(A|C)$ and $\mathbb{P}(B|C)$

$$\mathbb{P}(A|C) = \frac{\mathbb{P}(A \cap C)}{\mathbb{P}(C)} = \frac{\mathbb{P}(C|A).\mathbb{P}(A)}{\mathbb{P}(C)} = \frac{(0.75)(\frac{3}{5})}{0.81} = 0.5556$$

$$\mathbb{P}(B|C) = \frac{\mathbb{P}(B \cap C)}{\mathbb{P}(C)} = \frac{\mathbb{P}(C|B).\mathbb{P}(B)}{\mathbb{P}(C)} = \frac{(0.9)(\frac{2}{5})}{0.81} = 0.4444$$

Probability: Bayes's rule-Total Probability



- Let $A_1, \ A_2, \ A_3$ be a partition of Ω (i.e; $\Omega = A_1 \cup A_2 \cup A_3$ and $A_1 \cap A_2 \cap A_3 = \emptyset$).
- We know $\mathbb{P}(A_i) \rightarrow$ initial beliefs
- We know $\mathbb{P}(B|A_i)$, for every i $\rightarrow$ New information.
- One way of computing $\mathbb{P}(B)$

$$B = (B \cap A_1) \cup (B \cap A_2) \cup (B \cap A_3) \ \rightarrow \text{ These three sets are mutually exclusive}$$

$$\begin{aligned} \mathbb{P}(B) &= \mathbb{P}(B \cap A_1) + \mathbb{P}(B \cap A_2) + \mathbb{P}(B \cap A_3) \rightarrow \text{ Total probability} \\ &= \mathbb{P}(A_1)\mathbb{P}(B|A_1) + \mathbb{P}(A_2)\mathbb{P}(B|A_2) + \mathbb{P}(A_3)\mathbb{P}(B|A_3) \rightarrow \text{ conditional probability} \\ &= \sum_{i=1}^{3} \mathbb{P}(A_i)\mathbb{P}(B|A_i) \ \rightarrow \text{ Total probability} \rightarrow \text{ we can generalize it to n sets} \end{aligned}$$

- Wish to compute $\mathbb{P}(A_i|B) \rightarrow$ revise our beliefs given that B occurs

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(A_i \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A_i)\mathbb{P}(B|A_i)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A_i)\mathbb{P}(B|A_i)}{\sum_j \mathbb{P}(A_j)\mathbb{P}(B|A_j)}$$

Probability: Bayes's rule

- The multiplication rule:

$$\begin{aligned}\mathbb{P}(A \cap B \cap C) &= \mathbb{P}((A \cap B) \cap C) \\ &= \underline{\mathbb{P}(A \cap B)}\, \mathbb{P}(C | (A \cap B)) \\ &= \underbrace{\mathbb{P}(A)\, \mathbb{P}(B|A)}\, \mathbb{P}(C|(A \cap B))\end{aligned}$$

- We can generalize the rule to n events:

$$\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_n) = \mathbb{P}(A_1) \prod_{i=2}^{n} \mathbb{P}(A_i | A_1 \cap A_2 \cap \ldots \cap A_{i-1})$$

- For $n = 4$,

$$\begin{aligned}\mathbb{P}(A_1 \cap A_2 \cap A_3 \cap A_4) &= \mathbb{P}(A_1) \textstyle\prod_{i=2}^{4} \mathbb{P}(A_i | A_1 \cap A_2 \cap \ldots \cap A_{i-1}) \\ &= \mathbb{P}(A_1)\, \mathbb{P}(A_2|A_1)\, \mathbb{P}(A_3|A_1 \cap A_2)\, \mathbb{P}(A_4|A_1 \cap A_2 \cap A_3)\end{aligned}$$

Probability: Independence

Intuitively: Independence between two events stand for the fact that the first event, whether it occurred or not, doesn't give you any more information and does not cause you to change your beliefs about the second event.

$$\mathbb{P}(A|B) = \mathbb{P}(A) \text{ and } \mathbb{P}(B|A) = \mathbb{P}(B)$$

- If A and B are **independent**, so we have:

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

And

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

- Definition of independence: $\boxed{\mathbb{P}(A \cap B) = \mathbb{P}(A).\mathbb{P}(B)}$

Probability: Independence

Example 1:

1) Tossing a coin one time $\rightarrow |\Omega| = 2 \rightarrow \{H, T\} \rightarrow \mathbb{P}(H) = \mathbb{P}(T) = \dfrac{1}{2} = 0.5$

2) Tossing a coin two time

$$|\Omega| = 4 \rightarrow \{(H, T), (T, H), (H, H), (T, T)\} \rightarrow \mathbb{P}(H, H) = \frac{1}{4} = 0.25$$

$$\mathbb{P}(H, H) = \mathbb{P}(H).\mathbb{P}(H) = 0.5 \times 0.5 = 0.25$$

- Getting H at the 2^{nd} tossing is independent of getting H at the 1^{st} tossing

3) Tossing a coin three time

$$|\Omega| = 8 \rightarrow \{(H, H, H), (H, H, T), \ldots, (T, T)\} \rightarrow \mathbb{P}(H, H, H) = \frac{1}{8} = 0.125$$

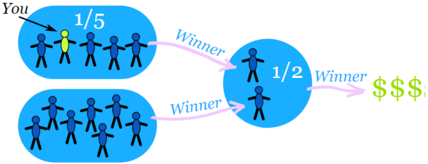
$$\mathbb{P}(H, H, H) = \mathbb{P}(H).\mathbb{P}(H).\mathbb{P}(H) = 0.5 \times 0.5 \times 0.5 = 0.125$$

- Getting H at the 3^{nd} tossing is independent of getting H at the 2^{nd} tossing and independent of getting H at the 1^{st} tossing

Probability: Independence

Example 2: There are two groups:

- A member of each group gets randomly chosen for the winners circle,
- Then one of those gets randomly chosen to get the big money prize



- What is your chance of winning the big prize?

▷ There is $\frac{1}{5}$ of chance to go to the winners circle and $\frac{1}{2}$ of chance to win the big prize.

→ So the probability of winning the big prize is $\frac{1}{5}$ followed by $\frac{1}{2}$ which makes:

$$\mathbb{P}(\text{winning the big prize}) = \frac{1}{5} \times \frac{1}{2} = \frac{1}{10} = 0.1$$

▷ Being selected at the first time and being selected the second time are two independent events.

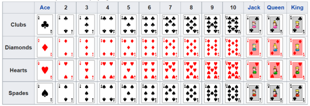
Probability: Independence

Example 3:

A card is chosen at random from a deck of 52 cards. Let these two different experiments:

1) The chosen card is replaced and a second card is chosen
2) The chosen card is not replaced and a second card is chosen

What is the probability of choosing a jack and then an eight for each case?



1) $\mathbb{P}(\text{jack}) = \frac{4}{52}, \quad \mathbb{P}(8) = \frac{4}{52}$

$$\mathbb{P}(\text{jack} \cap 8) = \mathbb{P}(\text{jack}).\mathbb{P}(8) = \frac{4}{52} \times \frac{4}{52} = \frac{16}{2704} = \frac{1}{169} \rightarrow \text{ independent events}$$

2) $\mathbb{P}(\text{jack}) = \frac{4}{52}, \quad \mathbb{P}(8) = \frac{4}{51}$

$$\mathbb{P}(\text{jack} \cap 8) = \mathbb{P}(\text{jack}).\mathbb{P}(8) = \frac{4}{52} \times \frac{4}{51} = \frac{16}{2652} \rightarrow \text{ independent events}$$

Probability: Independence

Example 4: Researchers surveyed recent graduates of two different universities about their annual incomes. The following two-way table displays data for the 300 graduates who responded to the survey (prior information).

Annual income	University A	University B	TOTAL
Under $20,000	36	24	60
$20,000 to 39,999	109	56	165
$40,000 and over	35	40	75
TOTAL	180	120	300

▷ Form this set of data we can compute the following probabilities

	University A	University B	Total
$< \$20K = E_1$	$\mathbb{P}(E_1 \cap A) = \frac{36}{300}$	$\mathbb{P}(E_1 \cap B) = \frac{24}{300}$	$\mathbb{P}(E_1) = \frac{36+24}{300} = \frac{60}{300}$
$\$20K - \$39.99 = E_2$	$\mathbb{P}(E_2 \cap A) = \frac{109}{300}$	$\mathbb{P}(E_2 \cap B) = \frac{56}{300}$	$\mathbb{P}(E_2) = \frac{109+56}{300} = \frac{165}{300}$
$\geq \$40K = E_3$	$\mathbb{P}(E_3 \cap A) = \frac{35}{300}$	$\mathbb{P}(E_3 \cap B) = \frac{40}{300}$	$\mathbb{P}(E_3) = \frac{35+40}{300} = \frac{75}{300}$
Total	$\mathbb{P}(A) = \frac{180}{300}$	$\mathbb{P}(B) = \frac{120}{300}$	1

Probability: Independence

	University A	University B	Total
$< \$20K = E_1$	$\mathbb{P}(E_1 \cap A) = \frac{36}{300}$	$\mathbb{P}(E_1 \cap B) = \frac{24}{300}$	$\mathbb{P}(E_1) = \frac{36+24}{300} = \frac{60}{300}$
$\$20K - \$39.99 = E_2$	$\mathbb{P}(E_2 \cap A) = \frac{109}{300}$	$\mathbb{P}(E_2 \cap B) = \frac{56}{300}$	$\mathbb{P}(E_2) = \frac{109+56}{300} = \frac{165}{300}$
$\geq \$40K = E_3$	$\mathbb{P}(E_3 \cap A) = \frac{35}{300}$	$\mathbb{P}(E_3 \cap B) = \frac{40}{300}$	$\mathbb{P}(E_3) = \frac{35+40}{300} = \frac{75}{300}$
Total	$\mathbb{P}(A) = \frac{180}{300}$	$\mathbb{P}(B) = \frac{120}{300}$	1

- Are the events "income is \$40 K and over (E_3)" and "attended University B" independent?

- **Method1:**

$$\mathbb{P}(E_3) = \frac{75}{300} = 0.25, \quad \mathbb{P}(E_3|B) = \frac{\mathbb{P}(E_3 \cap B)}{\mathbb{P}(B)} = \frac{40}{120} = 0.33$$

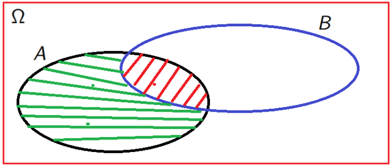
$\Rightarrow \ \mathbb{P}(E_3|B) \neq \mathbb{P}(E_3) \rightarrow$ They are **not independent**.

- **Method2:** $\mathbb{P}(E_3) = \frac{75}{300} = 0.25, \quad \mathbb{P}(B) = \frac{120}{300} = 0.4,$

$\mathbb{P}(E_3 \cap B) = \frac{40}{300} = 0.13 \neq \mathbb{P}(E_3).\mathbb{P}(B) = 0.1 \rightarrow$ They are **not independent**.

Probability: Independence

- **If A and B are independent, then A and B^c are also independent.**



$$A = (A \cap B) \cup (A \cap B^c) \;\rightarrow\; \text{and these two sets are disjoints}$$

$$\mathbb{P}(A) = \mathbb{P}[(A \cap B) \cup (A \cap B^c)] = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) = \underbrace{\mathbb{P}(A).\mathbb{P}(B)}_{\text{Independence}} + \mathbb{P}(A \cap B^c)$$

$$\Rightarrow \; \mathbb{P}(A \cap B^c) = \mathbb{P}(A) - \mathbb{P}(A)\,\mathbb{P}(B) = \mathbb{P}(A)[1 - \mathbb{P}(B)] = \mathbb{P}(A)\,\mathbb{P}(B^c)$$

$\rightarrow$ Which conclude the independence of A and B^c.

Probability: Independence

- **If A and B are independent, then A^c and B^c are also independent.**

$$\mathbb{P}(A^c \cap B^c) = \mathbb{P}[(A \cup B)^c] = 1 - \mathbb{P}(A \cap B) = 1 - [\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)]$$

$$= 1 - \mathbb{P}(A) - \mathbb{P}(B) + \underbrace{\mathbb{P}(A \cap B)}_{=\mathbb{P}(A).\mathbb{P}(B)} = 1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A).\mathbb{P}(B)$$

$$\Rightarrow \ \mathbb{P}(A^c \cap B^c) = [1 - \mathbb{P}(A)].[1 - \mathbb{P}(B)] = \mathbb{P}(A^c).\mathbb{P}(B^c)$$

$\rightarrow$ Which conclude the independence of A^c and B^c.

Probability: Independence

Independence of a collection of events:

- Events $A_1, A_2, \ldots, A_n$ are called **independents** if

$\mathbb{P}(A_i \cap A_j \cap \ldots A_m) = \mathbb{P}(A_i)\mathbb{P}(A_j)\ldots\mathbb{P}(A_m)$ for any **distinct indices** i,j, . . . , m

- For $n = 3$

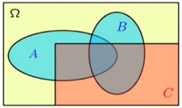
$$\left.\begin{array}{l} \bullet\ \mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2) \\ \bullet\ \mathbb{P}(A_1 \cap A_3) = \mathbb{P}(A_1)\mathbb{P}(A_3) \\ \bullet\ \mathbb{P}(A_2 \cap A_3) = \mathbb{P}(A_2)\mathbb{P}(A_3) \end{array}\right\} \Rightarrow \quad \text{pairwise independent}$$

- $\mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3)$

Probability: Conditional independence

▷ Two events A and B are **conditionally independent** given an event C with $\mathbb{P}(C) > 0$ if

$$\boxed{\mathrm{P}(\mathrm{A} \cap B|C) = \mathbb{P}(A|C)\ \mathbb{P}(B|C)}$$



Example:
A box contains two coins: a regular coin (C1) and one fake two-headed coin (C2) (i.e; $(\mathbb{P}(H) = 1)$). I choose a coin at random and toss it twice. Define the following events.

- A= First coin toss results in a H.
- B= Second coin toss results in a H.
- C1= regular coin has been selected.
- C2 = fake coin has been selected.

- Find $\mathbb{P}(A|C1), \mathbb{P}(B|C1), \mathbb{P}(A \cap B|C1), \mathbb{P}(A \cap B|C2), \mathbb{P}(A), \mathbb{P}(B)$, and $\mathbb{P}(A \cap B)$.

Probability: Conditional independence

Solution: We have the following information: $C1$:{H, T}, $C2$:{H}

$$\mathbb{P}(C1) = \frac{1}{2}, \quad \mathbb{P}(C2) = \frac{1}{2}, \quad \mathbb{P}(A|C1) = \frac{1}{2}, \quad \mathbb{P}(B|C1) = \frac{1}{2}, \quad \mathbb{P}(A|C2) = 1, \quad \mathbb{P}(B|C2) = 1$$

$$\mathbb{P}(A \cap B|C1) = \frac{1}{2}.\frac{1}{2} = \frac{1}{4}, \quad \mathbb{P}(A \cap B|C2) = 1$$

	regular coin (C1)	fake coin (C2)	
A	$\mathbb{P}(A \cap C1) = \mathbb{P}(A\|C1).\mathbb{P}(C1) = \frac{1}{4}$	$\mathbb{P}(A \cap C2) = \mathbb{P}(A\|C2).\mathbb{P}(C2) = \frac{1}{2}$	$\mathbb{P}(A) = \frac{3}{4}$
B	$\mathbb{P}(B \cap C1) = \mathbb{P}(B\|C1).\mathbb{P}(C1) = \frac{1}{4}$	$\mathbb{P}(B \cap C2) = \mathbb{P}(B\|C2).\mathbb{P}(C2) = \frac{1}{2}$	$\mathbb{P}(B) = \frac{3}{4}$

- $\mathbb{P}(A \cap B|C1) = \frac{1}{4} = \mathbb{P}(A|C1).\mathbb{P}(B|C1) = \frac{1}{2}.\frac{1}{2} = \frac{1}{4}$
- $\mathbb{P}(A \cap B|C2) = 1 = \mathbb{P}(A|C2).\mathbb{P}(B|C2) = 1 \times 1 = 1$

$\Rightarrow$ A and B are independent by knowing the chosen coin (conditionally independent)

- Let us compare $\mathbb{P}(A \cap B)$ and $\mathbb{P}(A).\mathbb{P}(B)$. We have

$$\mathbb{P}(A) = \mathbb{P}(A \cap C1) + \mathbb{P}(A \cap C2) = \frac{3}{4}, \quad \mathbb{P}(B) = \mathbb{P}(B \cap C1) + \mathbb{P}(B \cap C2) = \frac{3}{4}$$

Probability: Conditional independence

	regular coin (C1)	fake coin (C2)	
A	$\mathbb{P}(A \cap C1) = \mathbb{P}(A\|C1).\mathbb{P}(C1) = \frac{1}{4}$	$\mathbb{P}(A \cap C2) = \mathbb{P}(A\|C2).\mathbb{P}(C2) = \frac{1}{2}$	$\mathbb{P}(A) = \frac{3}{4}$
B	$\mathbb{P}(B \cap C1) = \mathbb{P}(B\|C1).\mathbb{P}(C1) = \frac{1}{4}$	$\mathbb{P}(B \cap C2) = \mathbb{P}(B\|C2).\mathbb{P}(C2) = \frac{1}{2}$	$\mathbb{P}(B) = \frac{3}{4}$

$$\begin{aligned}
\mathbb{P}(A \cap B) &= \mathbb{P}(A \cap B \cap C1) + \mathbb{P}(A \cap B \cap C2) \\
&= \mathbb{P}((A \cap B)|C1).\mathbb{P}(C1) + \mathbb{P}((A \cap B)|C2).\mathbb{P}(C2) \\
&= \frac{1}{4}.\frac{1}{2} + 1.\frac{1}{2} = \frac{5}{8}
\end{aligned}$$

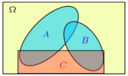
▷ As we see

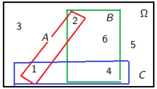
$$\mathbb{P}(A \cap B) = \frac{5}{8} \neq \mathbb{P}(A).\mathbb{P}(B) = \frac{3}{4}.\frac{3}{4} = \frac{9}{16}$$

⇒ A and B are **NOT independent (or dependent)** since $\mathbb{P}(A \cap B)$ depend on the chosen coin, but they are **conditionally independent knowing in advance which coin is tossed**.

Probability: Conditional independence

- **Assume** A **and** B **are independent**. If we told that C occurred, **are** A **and** B **still independent?** $\rightarrow$ we can have two events that they are **independent** but **not conditionally independent** given an event C.





Example: Consider the rolling of a die $\rightarrow \Omega = \{1,2,3,4,5,6\}$. Let the following events

$A = \{1,2\},\ B = \{2,4,6\},\ C = \{1,4\} \rightarrow \mathbb{P}(A) = \frac{2}{6} = \frac{1}{3},\ \mathbb{P}(B) = \frac{3}{6} = \frac{1}{2},\ \mathbb{P}(C) = \frac{2}{6} = \frac{1}{3}$

$$\mathbb{P}(A \cap B) = \mathbb{P}(\{2\}) = \frac{1}{6} = \mathbb{P}(A).\mathbb{P}(B) = \frac{1}{3}.\frac{1}{2} = \frac{1}{6} \rightarrow \text{ A and B are } \textbf{independent}$$

$$\mathbb{P}(A \cap B|C) = \frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(C)} = \frac{\mathbb{P}(\emptyset)}{\mathbb{P}(C)} = 0$$

$$\mathbb{P}(A|C) = \frac{\mathbb{P}(A \cap C)}{\mathbb{P}(C)} = \frac{|A \cap C|}{|C|} = \frac{1}{2} \neq 0 \ \text{ and } \ \mathbb{P}(B|C) = \frac{\mathbb{P}(B \cap C)}{\mathbb{P}(C)} = \frac{|B \cap C|}{|C|} = \frac{1}{2} \neq 0$$

$\Rightarrow \mathbb{P}(A \cap B|C) \neq \mathbb{P}(A|C)\ \mathbb{P}(B|C) \Rightarrow$ They are **not conditionally independent**

Probability: Conditional independence

▷ A and B are **conditionally independent** knowing C, means that if a given knowledge that C occurs, so A and B becomes independent (i.e; knowledge of whether A occurs provides no information on the likelihood of B occurring, and knowledge of whether B occurs provides no information on the likelihood of A occurring).

Probability: Conditional independence

Properties: Let the following properties

▷ Suppose that A and B are conditionally independent knowing that the event C occurs:

$$\mathbb{P}(A\cap B|C) = \frac{\mathbb{P}(A\cap B\cap C)}{\mathbb{P}(C)} = \underbrace{\mathbb{P}(A|C).\mathbb{P}(B|C)}_{\text{conditional independence}} .$$

▷ We can deduce the following properties:

1) $\mathbb{P}(A\cap B^c|C) = \mathbb{P}(A|C).\mathbb{P}(B^c|C)$

- We have $\mathbb{P}(A\cap B^c|C) = \dfrac{\mathbb{P}(A\cap B^c\cap C)}{\mathbb{P}(C)}$. On the other hand

$$\mathbb{P}(A\cap C) = \mathbb{P}(A\cap C\cap B) + \mathbb{P}(A\cap C\cap B^c) \Rightarrow \frac{\mathbb{P}(A\cap C)}{\mathbb{P}(C)} = \frac{\mathbb{P}(A\cap C\cap B)}{\mathbb{P}(C)} + \frac{\mathbb{P}(A\cap C\cap B^c)}{\mathbb{P}(C)}$$

$$\Rightarrow \mathbb{P}(A|C) = \mathbb{P}(A\cap B|C) + \mathbb{P}(A\cap B^c|C) \Rightarrow \mathbb{P}(A|C) = \mathbb{P}(A|C).\mathbb{P}(B|C) + \mathbb{P}(A\cap B^c|C)$$

$$\Rightarrow \mathbb{P}(A\cap B^c|C) = \mathbb{P}(A|C) - \mathbb{P}(A|C).\mathbb{P}(B|C) = \mathbb{P}(A|C)[1-\mathbb{P}(B|C)] = \mathbb{P}(A|C).\mathbb{P}(B^c|C)$$

Probability: Conditional independence

2) $\mathbb{P}(A^c \cap B|C) = \mathbb{P}(A^c|C).\mathbb{P}(B|C)$

$$\mathbb{P}(A^c \cap B|C) = \frac{\mathbb{P}(A^c \cap B \cap C)}{\mathbb{P}(C)}$$

On the other hand

$$\mathbb{P}(B \cap C) = \mathbb{P}(B \cap C \cap A) + \mathbb{P}(B \cap C \cap A^c) \Rightarrow \frac{\mathbb{P}(B \cap C)}{\mathbb{P}(C)} = \frac{\mathbb{P}(B \cap C \cap A)}{\mathbb{P}(C)} + \frac{\mathbb{P}(B \cap C \cap A^c)}{\mathbb{P}(C)}$$

$$\Rightarrow \mathbb{P}(B|C) = \mathbb{P}(A \cap B|C) + \mathbb{P}(A^c \cap B|C) \Rightarrow \mathbb{P}(B|C) = \mathbb{P}(A|C).\mathbb{P}(B|C) + \mathbb{P}(A^c \cap B|C)$$

$$\Rightarrow \mathbb{P}(A^c \cap B|C) = \mathbb{P}(B|C) - \mathbb{P}(A|C).\mathbb{P}(B|C) = \mathbb{P}(B|C)[1 - \mathbb{P}(A|C)] = \mathbb{P}(B|C).\mathbb{P}(A^c|C)$$

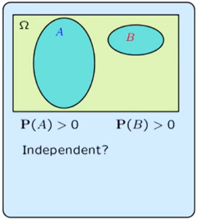
Probability: Conditional independence

3) $\mathbb{P}(A^c \cap B^c|C) = \mathbb{P}(A^c|C).\mathbb{P}(B^c|C)$

$$
\begin{aligned}
\mathbb{P}(A^c \cap B^c|C) &= \mathbb{P}((A \cup B)^c|C) \\
&= 1 - \mathbb{P}((A \cup B)^c|C) \\
&= 1 - \frac{\mathbb{P}((A \cup B) \cap C)}{\mathbb{P}(C)} \\
&= 1 - \frac{\mathbb{P}((A \cap C) \cup (B \cap C))}{\mathbb{P}(C)} \\
&= 1 - \frac{\mathbb{P}(A \cap C) + \mathbb{P}(B \cap C) - \mathbb{P}(A \cap B \cap C)}{\mathbb{P}(C)} \\
&= 1 - \frac{\mathbb{P}(A \cap C)}{\mathbb{P}(C)} - \frac{\mathbb{P}(B \cap C)}{\mathbb{P}(C)} + \frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(C)} \\
&= 1 - \mathbb{P}(A|C) - \mathbb{P}(B|C) + \mathbb{P}(A \cap B|C) \\
&= 1 - \mathbb{P}(A|C) - \mathbb{P}(B|C) + \mathbb{P}(A|C).\mathbb{P}(B|C) \\
&= [1 - \mathbb{P}(A|C)].[1 - \mathbb{P}(B|C)] \\
&= \mathbb{P}(A^c|C).\mathbb{P}(B^c|C)
\end{aligned}
$$

Probability: Independence

- Don't confuse independence and disjoints:

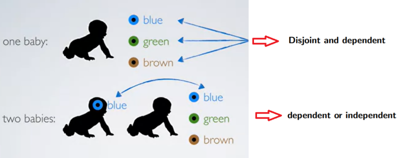


• $A \cap B = \emptyset \;\Rightarrow\; \mathbb{P}(A \cap B) = 0 \Rightarrow \mathbb{P}(A|B) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = 0 \rightarrow$ It is an impossible event $\rightarrow$ But $\mathbb{P}(A) \neq 0$ and $\mathbb{P}(B) \neq 0 \Rightarrow \mathbb{P}(A).\mathbb{P}(B) \neq 0 \Rightarrow \mathbb{P}(A \cap B) \neq \mathbb{P}(A).\mathbb{P}(B) \rightarrow$ These events are **disjoints** but **dependent**

• If A and B are **independent** they should not be **disjoint**

• $\mathbb{P}(\emptyset|B) = \dfrac{\mathbb{P}(\emptyset \cap B)}{\mathbb{P}(B)} = \dfrac{\mathbb{P}(\emptyset)}{\mathbb{P}(B)} = \dfrac{0}{\mathbb{P}(B)} = 0 = \mathbb{P}(\emptyset) \rightarrow$ **The impossible event is independent to all events** and **disjoint with all events**

Probability: Independence



• Suppose that we have one baby (B_1) and consider the colors {Blue, Green, Brown}.
- Event B: The baby has blue eyes $\rightarrow \mathbb{P}(B) = \frac{1}{3}$
- Event G: The baby has green eyes $\rightarrow \mathbb{P}(G) = \frac{1}{3}$
- Event R: The baby has brown eyes $\rightarrow \mathbb{P}(R) = \frac{1}{3}$

$B \cap G \cap R = \emptyset \rightarrow$ the baby cannot has all these colors at the same time $\rightarrow$ **disjoints**

$\mathbb{P}(G|B) = 0 \neq \mathbb{P}(G)$ $\rightarrow$ **dependent** $\Rightarrow$ The occurrence of the event B will affect the occurrence of the event G. $\rightarrow$ if we know that the baby has blue eyes, so G and R cannot happen.

• Suppose that we have two babies (B_1) and (B_2) $\rightarrow$ The color of the eyes of the two babies are **independent** (can be dependent if we will add more information) but they are **not disjoints** $\rightarrow$ the two babies can have the same color of eyes.

Probability: Independence

- Independent vs disjoint events:

▷ Events are considered **disjoint** if they never occur at the same time. Events are considered **independent** if they are unrelated.

▷ **Example1:** Flipping a Coin

- **Scenario 1:** Suppose we flip a coin once $\rightarrow \Omega = \{H, T\}$. Let the events:

A: The coin landing on head $= \{H\} \rightarrow \ \mathbb{P}(A) = \frac{1}{2}$

B: The coin landing on tail $= \{T\} \rightarrow \ \mathbb{P}(B) = \frac{1}{2}$

• Event A and event B are disjoint $(A \cap B = \emptyset) \rightarrow$ the coin can't possibly land on heads and tails at the same time, but they are dependent since $\mathbb{P}(A|B) = 0 \neq \mathbb{P}(A).\mathbb{P}(B)$.

- **Scenario 2:** Suppose we flip a coin twice $\rightarrow \Omega = \{HH, TH, HT, TT\}$. Let the events

A: The coin landing on head on the first flip $= \{HH, HT\} \rightarrow \ \mathbb{P}(A) = \frac{2}{4} = \frac{1}{2}$

B: The coin landing on head on the second flip $= \{HH, TH\} \rightarrow \ \mathbb{P}(B) = \frac{2}{4} = \frac{1}{2}$

• Event A and event B are not disjoints and they are independent because the outcome of one coin flip doesn't affect the outcome of the other.

$$\mathbb{P}(A \cap B) = \mathbb{P}(\{HH\}) = \frac{1}{4} = \mathbb{P}(A).\mathbb{P}(B) = \frac{1}{2}.\frac{1}{2} \quad \text{or} \quad \mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2} = \mathbb{P}(A)$$

Probability: Independence

▷ **Example2:** Rolling a Dice

Scenario 1: Suppose we roll a dice once $\rightarrow \Omega = \{1, 2, 3, 4, 5, 6\}$. Let

A: The dice lands on an even number = {2,4,6} $\rightarrow \mathbb{P}(A) = \frac{3}{6} = \frac{1}{2}$

B: The dice lands on an odd number = {1,3,5} $\rightarrow \mathbb{P}(B) = \frac{3}{6} = \frac{1}{2}$

• Event A and event B are **disjoint** because the dice can't possibly land on an even number and an odd number at the same time but they are dependent since if A occurs B cannot occur $\rightarrow \mathbb{P}(A|B) = 0 \neq \mathbb{P}(A).\mathbb{P}(B)$.

- **Scenario 2:** Suppose we roll a dice twice. Let

A: The dice lands on a "5" on the first roll

B: The dice lands on a "5" on the second roll

• Event A and event B are not disjoints but they are **independent** because the outcome of one dice roll doesn't affect the outcome of the other.

$$\mathbb{P}(A) = \frac{6}{36} = \frac{1}{6}, \quad \mathbb{P}(B) = \frac{6}{36} = \frac{1}{6}, \quad \mathbb{P}(A \cap B) = \mathbb{P}(5, 5) = \frac{1}{36} = \frac{1}{6}.\frac{1}{6} = \mathbb{P}(A).\mathbb{P}(B)$$

Examples- Training

Example 1: We throw 2 dice $\rightarrow \Omega = \{(1,1),(1,2),\ldots,(6,6)\} \rightarrow$ 36 elements. Compute

1) $\mathbb{P}(\text{sum of 2 faces is 9})$
2) $\mathbb{P}(\text{sum of 2 faces is 9}|\text{the first face is 4})$

Solution:

1) Without prior information:

$$\mathbb{P}(\text{sum of 2 faces is 9}) = \mathbb{P}(\{(3,6),(6,3),(4,5),(5,4)\}) = \frac{4}{36} = \frac{1}{9}$$

2) With additional information: If first face is 4. Then

$$\mathbb{P}(\text{first face is 4}) = \mathbb{P}(\{(4,1),(4,2),(4,3),(4,4),(4,5),(4,6)\}) = \frac{6}{36} = \frac{1}{6}$$

$$\mathbb{P}(\text{sum of 2 faces is 9}|\text{the first face is 4}) = \frac{\mathbb{P}(\{\text{sum of 2 faces is 9}\} \cap \{\text{the first face is 4}\})}{\mathbb{P}(\text{first face is 4})}$$

$$\frac{\mathbb{P}[(4,5)]}{\frac{1}{6}} = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{6}$$

- With additional information, probability of having sum$=9$ becomes $\frac{1}{6}$.

Examples- Training

Example 2:

An urn contain 8 red and 4 white balls. We draw 2 balls **without replacement**. Let

R1 = 1st ball drawn is red

R2 = 2nd ball drawn is red

- Find $\mathbb{P}(R1 \cap R2)$
- Find $\mathbb{P}(R2)$

Solution: Information that we can collect are:

$$\mathbb{P}(R1) = \frac{8}{12} = \frac{2}{3}, \quad \mathbb{P}(R1^c) = \frac{4}{12}, \quad \mathbb{P}(R2|R1) = \frac{7}{11}, \quad \mathbb{P}(R2|R1^c) = \frac{8}{11}.$$

$$\mathbb{P}(R1 \cap R2) = \mathbb{P}(R2|R1).\mathbb{P}(R1) = \frac{7}{11}.\frac{2}{3} = \frac{14}{33} = 0.42$$

$$\mathbb{P}(R2) = \mathbb{P}(R2 \cap R1) + \mathbb{P}(R2 \cap R1^c) = \underbrace{\mathbb{P}(R1).\mathbb{P}(R2|R1)}_{\frac{8}{12}.\frac{7}{11}} + \underbrace{\mathbb{P}(R1^c).\mathbb{P}(R2|R1^c)}_{\frac{4}{12}.\frac{8}{11}} = \frac{88}{132} = 0.66$$

- Note that $\mathbb{P}(R1 \cap R2) = 0.42 \neq \mathbb{P}(R1).\mathbb{P}(R2) = 0.44 \rightarrow$ These events are **dependent**

Examples- Training

- ▷ We can construct the following table:

$$\mathbb{P}(R1) = \frac{2}{3},\ \mathbb{P}(R_1^c) = \frac{1}{3},\ \mathbb{P}(R2|R1) = \frac{7}{11},\quad \mathbb{P}(R2|R1^c) = \frac{8}{11},\ \mathbb{P}(R_2^c|R1) = \frac{4}{11},\ \mathbb{P}(R_2^c|R_1^c) = \frac{3}{11}$$

	R1	R_1^c	Total
R2	$\mathbb{P}(R1 \cap R2) = \mathbb{P}(R2|R1).\mathbb{P}(R1) = \frac{14}{33}$	$\mathbb{P}(R_1^c \cap R2) = \mathbb{P}(R2|R_1^c).\mathbb{P}(R_1^c) = \frac{8}{33}$	$\frac{2}{3}$
R_2^c	$\mathbb{P}(R1 \cap R_2^c) = \mathbb{P}(R_2^c|R1).\mathbb{P}(R1) = \frac{8}{33}$	$\mathbb{P}(R_1^c \cap R_2^c) = \mathbb{P}(R_2^c|R_1^c).\mathbb{P}(R_1^c) = \frac{3}{33}$	$\frac{1}{3}$
Total	$\mathbb{P}(R_1) = \frac{2}{3}$	$\mathbb{P}(R_1^c) = \frac{1}{3}$	1

$$\mathbb{P}(R1 \cap R2) = \frac{14}{33} = 0.42$$

$$\mathbb{P}(R2) = \mathbb{P}(R2 \cap R1) + \mathbb{P}(R2 \cap R1^c) = \frac{2}{3} = 0.66,\quad \mathbb{P}(R1) = \frac{2}{3}$$

$$\mathbb{P}(R1 \cap R2) = \frac{14}{33} \neq \mathbb{P}(R1).\mathbb{P}(R2) = \frac{2}{3}.\frac{2}{3} = \frac{4}{9}$$

Examples- Training

Example 3:

An urn contain 8 Red and 4 White balls. We draw 2 balls **with replacement**. Let

R1 = 1st ball drawn is red

R2 = 2nd ball drawn is red

- Find $\mathbb{P}(R1 \cap R2)$
- Find $\mathbb{P}(R2)$

Solution: Information that we can collect are:

$$\mathbb{P}(R1) = \frac{8}{12} = \frac{2}{3}, \quad \mathbb{P}(R2|R1) = \frac{8}{12} = \frac{2}{3}, \quad \mathbb{P}(R1^c) = \frac{4}{12}, \quad \mathbb{P}(R2|R1^c) = \frac{8}{12}.$$

$$\mathbb{P}(R1 \cap R2) = \mathbb{P}(R1).\mathbb{P}(R2|R1) = \frac{8}{12}.\frac{8}{12} = \frac{2}{3}.\frac{2}{3} = \frac{4}{9} = 0.44$$

$$\mathbb{P}(R2) = \mathbb{P}(R2 \cap R1) + \mathbb{P}(R2 \cap R1^c) = \underbrace{\mathbb{P}(R1).\mathbb{P}(R2|R1)}_{\frac{8}{12}.\frac{8}{12}} + \underbrace{\mathbb{P}(R1^c).\mathbb{P}(R2|R1^c)}_{\frac{4}{12}.\frac{8}{12}} = \frac{2}{3} = 0.66$$

- Note that $\mathbb{P}(R1 \cap R2) = 0.44 = \mathbb{P}(R1).\mathbb{P}(R2) = \frac{2}{3}.\frac{2}{3} = 0.44 \rightarrow$ These events **are independent.**

Examples- Training

▷ We can construct the following table:

$$\mathbb{P}(R1) = \frac{2}{3},\ \mathbb{P}(R_1^c) = \frac{1}{3},\ \mathbb{P}(R2|R1) = \frac{2}{3},\ \mathbb{P}(R2|R1^c) = \frac{2}{3},\ \mathbb{P}(R_2^c|R1) = \frac{1}{3},\ \mathbb{P}(R_2^c|R_1^c) = \frac{1}{3}$$

	R1	R_1^c	Total
R2	$\mathbb{P}(R1 \cap R2) = \mathbb{P}(R2\|R1).\mathbb{P}(R1) = \frac{4}{9}$	$\mathbb{P}(R_1^c \cap R2) = \mathbb{P}(R2\|R_1^c).\mathbb{P}(R_1^c) = \frac{2}{9}$	$\frac{2}{3}$
R_2^c	$\mathbb{P}(R1 \cap R_2^c) = \mathbb{P}(R_2^c\|R1).\mathbb{P}(R1) = \frac{2}{9}$	$\mathbb{P}(R_1^c \cap R_2^c) = \mathbb{P}(R_2^c\|R_1^c).\mathbb{P}(R_1^c) = \frac{1}{9}$	$\frac{1}{3}$
Total	$\mathbb{P}(R_1) = \frac{2}{3}$	$\mathbb{P}(R_1^c) = \frac{1}{3}$	1

$$\mathbb{P}(R1 \cap R2) = \frac{4}{9} = 0.44$$

$$\mathbb{P}(R2) = \mathbb{P}(R2 \cap R1) + \mathbb{P}(R2 \cap R1^c) = \frac{2}{3}, \quad \mathbb{P}(R1) = \frac{2}{3}$$

$$\mathbb{P}(R1 \cap R2) = \frac{4}{9} = \mathbb{P}(R1).\mathbb{P}(R2) = \frac{2}{3}.\frac{2}{3} = \frac{4}{9}$$

Examples- Training

Example 4:

An insurance company divides its customers into three classes R_1, R_2 and R_3 : good risks, medium risks, and bad risks, respectively. The numbers of these three classes represent 20% of the total population for the class R_1, 50% for the class R_2 and 30% for the class R_3. Statistics indicate that the probabilities of having an accident during the year for a person in one of these three classes are respectively 0.05, 0.15 and 0.30.

1) What is the probability that a randomly selected person from the population has an accident during the year?
2) If Mr. Martin has not had an accident this year, what is the likelihood that he is a good risk?
3) Are the class of risks independent on having an accident during the year?

Examples- Training

Example 4- Solution- (Method 1)

- Event R_1: good risk $\to \mathbb{P}(R_1) = 0.2$
- Event R_2: medium risk $\to \mathbb{P}(R_2) = 0.5$
- Event R_3: bad risk $\to \mathbb{P}(R_1) = 0.3$
- Event A: "having an accident".

$$\mathbb{P}(A|R_1) = 0.05 = 5\%, \quad \mathbb{P}(A|R_2) = 0.15 = 15\%, \quad \mathbb{P}(A|R_3) = 0.3 = 30\%$$

	Have an accident (A)	Don't have an accident (A^c)	Total
R1	$\mathbb{P}(A \cap R1) = \mathbb{P}(A\|R1).\mathbb{P}(R1) = 0.01$	$\mathbb{P}(A^c \cap R1) = \mathbb{P}(A^c\|R1).\mathbb{P}(R1) = 0.19$	0.2
R2	$\mathbb{P}(A \cap R2) = \mathbb{P}(A\|R2).\mathbb{P}(R2) = 0.075$	$\mathbb{P}(A^c \cap R2) = \mathbb{P}(A^c\|R2).\mathbb{P}(R1) = 0.425$	0.5
R3	$\mathbb{P}(A \cap R3) = \mathbb{P}(A\|R3).\mathbb{P}(R3) = 0.09$	$\mathbb{P}(A^c \cap R3) = \mathbb{P}(A^c\|R3).\mathbb{P}(R3) = 0.21$	0.3
Total	0.175	0.825	1

1) $\mathbb{P}(A) = 0.175$

2) $\mathbb{P}(R1|A^c) = \dfrac{\mathbb{P}(R1 \cap A^c)}{\mathbb{P}(A^c)} = \dfrac{0.19}{0.825} = 0.2303$

3) The class of risks dependent on having an accident during the year, since

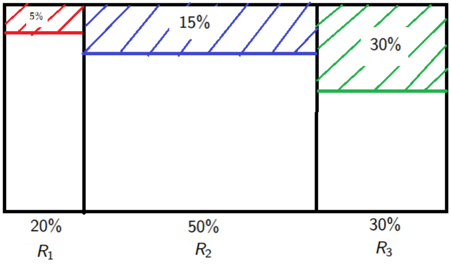
$$\mathbb{P}(A|R_1) = 0.05 \neq \mathbb{P}(A) = 0.175, \quad \mathbb{P}(A|R_2) = 0.15 \neq \mathbb{P}(A), \quad \mathbb{P}(A|R_3) = 0.3 \neq \mathbb{P}(A)$$

Examples- Training

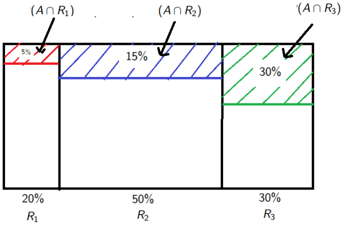
Example 4- Solution (Method 2)

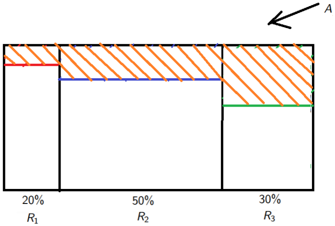
- Event R_1: good risk $\rightarrow \mathbb{P}(R_1) = 0.2$
- Event R_2: medium risk $\rightarrow \mathbb{P}(R_2) = 0.5$
- Event R_3: bad risk $\rightarrow \mathbb{P}(R_1) = 0.3$
- Event A: "having an accident".

$$\mathbb{P}(A|R_1) = 0.05 = 5\%, \quad \mathbb{P}(A|R_2) = 0.15 = 15\%, \quad \mathbb{P}(A|R_3) = 0.3 = 30\%$$



Examples- Training

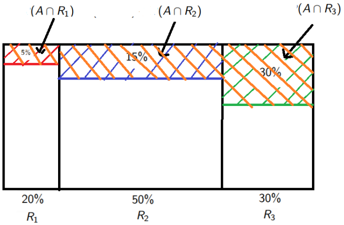


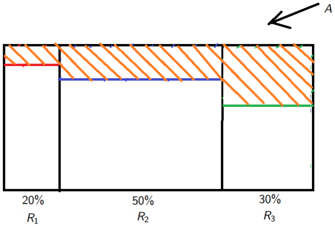


1) What is the probability that a randomly selected person from the population has an accident during the year?

$$\begin{aligned}
\mathbb{P}(A) &= \mathbb{P}(A \cap R_1) + \mathbb{P}(A \cap R_2) + \mathbb{P}(A \cap R_3) \\
&= \mathbb{P}(A|R_1)\mathbb{P}(R_1) + \mathbb{P}(A|R_2)\mathbb{P}(R_2) + \mathbb{P}(A|R_3)\mathbb{P}(R_3) \\
&= (0.05 \times 0.2) + (0.15 \times 0.5) + (0.3 \times 0.3) \\
&= 0.01 + 0.075 + 0.09 = 0.175
\end{aligned}$$

Examples- Training





2) If Mr. Martin has not had an accident this year, what is the likelihood that he is a good risk?

$$\mathbb{P}(R_1|A^c) = \frac{\mathbb{P}(R_1 \cap A^c)}{\mathbb{P}(A^c)} = \frac{\mathbb{P}(A^c|R_1)\mathbb{P}(R_1)}{\mathbb{P}(A^c)} = \frac{(1-\mathbb{P}(A|R_1))\mathbb{P}(R_1)}{1-\mathbb{P}(A)}$$

$$= \frac{(1-0.05)\ 0.2}{1-0.175} = \frac{0.19}{0.825} = 0.2303$$

3) The class of risks dependent on having an accident during the year, since

$$\mathbb{P}(A|R_1) = 0.05 \neq \mathbb{P}(A) = 0.175, \quad \mathbb{P}(A|R_2) = 0.15 \neq \mathbb{P}(A), \quad \mathbb{P}(A|R_3) = 0.3 \neq \mathbb{P}(A)$$

Examples- Training

Example 5:

In the teachers' room 60% are women and 40% are men; one in three women wears glasses and one in two men wears glasses:

- what is the probability that a random eyeglass wearer is a woman?

Example 5-Solution (Method 1):

- Event W:{The teacher is a woman} $\rightarrow$ $\mathbb{P}(W) = 0.6$
- Event M:{The teacher is a man} $\rightarrow$ $\mathbb{P}(M) = 0.4$
- Event G:{wears glasses} $\rightarrow$ $\mathbb{P}(G|W) = \frac{1}{3}$, $\mathbb{P}(G|M) = \frac{1}{2}$

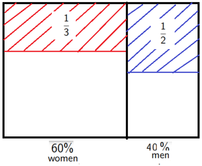
	Wears glasses (G)	Don't wear glasses (G^c)	Total
Woman (W)	$\mathbb{P}(W \cap G) = \mathbb{P}(G\|W).\mathbb{P}(W) = 0.2$	$\mathbb{P}(W \cap G^c) = 0.6 - 0.2 = 0.4$	0.6
Man	$\mathbb{P}(M \cap G) = \mathbb{P}(G\|M).\mathbb{P}(M) = 0.2$	$\mathbb{P}(M \cap G^c) = 0.4 - 0.2 = 0.2$	0.4
Total	0.4	0.6	1

- What is the probability that a random eyeglass wearer is a woman?

$$\mathbb{P}(W|G) = \frac{\mathbb{P}(W \cap G)}{\mathbb{P}(G)} = \frac{0.2}{0.4} = 0.5$$

Examples- Training

Example 5-Solution (Method 2):



- What is the probability that a random eyeglass wearer is a woman?

$$\mathbb{P}(W|G) = \frac{\mathbb{P}(W \cap G)}{\mathbb{P}(G)} = \frac{\mathbb{P}(G|W)\mathbb{P}(W)}{\mathbb{P}(G)}$$

- We need to compute $\mathbb{P}(G)$:

$$\Rightarrow \mathbb{P}(G) = \mathbb{P}(G \cap W) + \mathbb{P}(G \cap M) = \mathbb{P}(G|W)\mathbb{P}(W) + \mathbb{P}(G|M)\mathbb{P}(M) = \frac{0.6}{3} + \frac{0.4}{2} = 0.2 + 0.2 = 0.4$$

$$\Rightarrow \mathbb{P}(W|G) = \frac{\mathbb{P}(G|W)\mathbb{P}(W)}{\mathbb{P}(G)} = \frac{0.2}{0.4} = 0.5$$

Examples- Training

Example 6: Weather Forecasting

One of the most common real life examples of using conditional probability is weather forecasting.

▷ Suppose the following two probabilities are known:

- $\mathbb{P}(\text{cloudy}) = 0.25$
- $\mathbb{P}(\text{rainy} \cap \text{cloudy}) = 0.15$

$$\mathbb{P}(\text{rainy}|\text{cloudy}) = \frac{\mathbb{P}(\text{rainy} \cap \text{cloudy})}{\mathbb{P}(\text{cloudy})} = \frac{0.15}{0.25} = 0.6 \rightarrow 60\%$$

Examples- Training

Example 7: Sports Betting

Conditional probability is frequently used by sports betting companies to determine the odds they should set for certain teams to win certain games.

▷ Suppose the following two probabilities are known about some basketball team:

- $\mathbb{P}(\text{Team A star player is hurt}) = 0.15$
- $\mathbb{P}(\text{Team A wins} \cap \text{Team A start player is hurt}) = 0.02$

$$\mathbb{P}(\text{Team A wins}|\text{Team A start player is hurt}) =$$

$$\frac{\mathbb{P}(\text{Team A wins} \cap \text{Team A start player is hurt})}{\mathbb{P}(\text{Team A start player is hurt})} = \frac{0.02}{0.15} = 0.13 \rightarrow 13\%$$

Examples- Training

Example 8: Suppose that we have the following prior information

	Have pets (P)	Do not have pets (P^c)	Total
Male (M)	0.41	0.08	0.49
Female (F)	0.45	0.06	0.51
Total	0.86	0.14	1

- What is the probability that a randomly selected person is male, knowing that he own a pet? (i.e; $\mathbb{P}(M|P) = \dfrac{\mathbb{P}(M \cap P)}{\mathbb{P}(P)}$ =?)

- $\mathbb{P}(M \cap P) = 0.41$
- $\mathbb{P}(P) = \mathbb{P}(P \cap M) + \mathbb{P}(P \cap F) = 0.41 + 0.45 = 0.86$
- $\mathbb{P}(M|P) = \dfrac{\mathbb{P}(M \cap P)}{\mathbb{P}(P)} = \dfrac{0.41}{0.86} = 0.4767$
- $\mathbb{P}(M|P) = 0.4767 \neq \mathbb{P}(M) = 0.49 \ \rightarrow$ M and P are not independent

Examples- Training

▷ Reliability:

Real-life systems often are composed of several components. For example, a system may consist of two components that are connected in **parallel** as shown in Figure 1, or in **series** as shown in Figure 2.

• **Parallel connection:** When the system's components are connected in parallel, the system works if **at least one** of the components is functional.

• **Series connection:** When the system's components are connected in series, the system works if **all of the components** are functional.

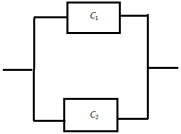


Figure 1. Parallel

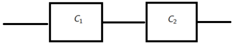


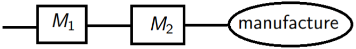
Figure 2. Series

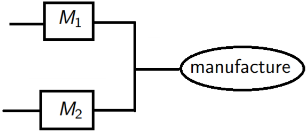
- Let the event S: "The system is functional"
- Parallel case → $S = \{C_1 \text{ OR } C_2 \text{ are functional}\} \Rightarrow S = C_1 \cup C_2$.
- Series case → $S = \{C_1 \text{ AND } C_2 \text{ are functional}\} \Rightarrow S = C_1 \cap C_2$.

Examples- Training

Example 9:

In a factory, two machines M_1 and M_2 are used jointly to manufacture cylindrical parts. These machines are connected in series or in parallel.





▷ For a given period, their probabilities of breaking down are respectively 0.01 and 0.008. Moreover, the probability of the event "the machine M_2 is down knowing that M_1 is down" is equal to 0.4.

1. What is the probability of having both machines down at the same time?
2. What is the probability that the manufacture in parallel case is working?
3. What is the probability that the manufacture in series case is working?

Examples- Training

Example 9-Solution

- Event A_1: $\{\text{Machine } M_1 \text{ is breaking down}\} \rightarrow \mathbb{P}(A_1) = 0.01$
- Event A_2: $\{\text{Machine } M_2 \text{ is breaking down}\} \rightarrow \mathbb{P}(A_2) = 0.008$
- The probability of the event "the machine M_2 is down knowing that M_1 is down" is equal to 0.4 $\Rightarrow \mathbb{P}(A_2|A_1) = 0.4$.

1. What is the probability of having both machines down at the same time?

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_2|A_1)\,\mathbb{P}(A_1) = (0.01)(0.4) = 0.004$$

2. What is the probability that manufacture in parallel case is working?

$$\mathbb{P}(A_1^c \cup A_2^c) = \mathbb{P}((A_1 \cap A_2)^c) = 1 - \mathbb{P}(A_1 \cap A_2) = 1 - 0.004 = 0.996.$$

3. What is the probability that the manufacture in series case is working?

$$\mathbb{P}(A_1^c \cap A_2^c) = \mathbb{P}(A_1^c) + \mathbb{P}(A_2^c) - \mathbb{P}(A_1^c \cup A_2^c)$$
$$= (1 - 0.01) + (1 - 0.008) - 0.996 = 0.986$$

Examples- Training

▷ We can construct the following table according to the given information:

$$\mathbb{P}(A_1) = 0.01,\ \mathbb{P}(A_2) = 0.008,\ \mathbb{P}(A_2|A_1) = 0.4$$

$$\mathbb{P}(A_1^c \cap A_2) = 0.008 - 0.004 = 0.004$$

$$\mathbb{P}(A_2^c \cap A_1) = \mathbb{P}(A_2^c|A_1).\mathbb{P}(A_1) = [1 - \mathbb{P}(A_2|A_1)].\mathbb{P}(A_1) = 0.6 \times 0.01 = 0.006$$

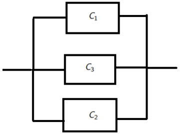
$$\mathbb{P}(A_1^c \cap A_2^c) = 0.992 - 0.006 = 0.992$$

	A_1	A_1^c	Total
A_2	$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_2\|A_1).\mathbb{P}(A_1) = 0.004$	$\mathbb{P}(A_1^c \cap A_2) = 0.004$	0.008
A_2^c	$\mathbb{P}(A_2^c \cap A_1) = \mathbb{P}(A_2^c\|A_1).\mathbb{P}(A_1) = 0.006$	$\mathbb{P}(A_1^c \cap A_2^c) = 0.986$	0.992
Total	0.01	0.99	1

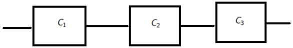
Examples- Training

Example 10: Consider the following systems and assume that component k is functional with probability P_k and it is independent to other components.

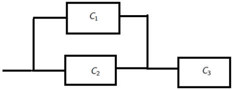
- Compute the probability that the system is functional in each of the following cases:



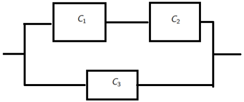
1



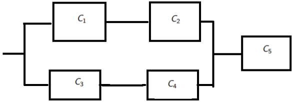
2



3



4



5

(1) $S = C_1 \cup C_2 \cup C_3$

(2) $S = C_1 \cap C_2 \cap C_3$

(3) $S = C_3 \cap (C_1 \cup C_2)$

(4) $S = (C_1 \cap C_2) \cup C_3$

(5) $S = C_5 \cap [(C_1 \cap C_2) \cup (C_3 \cap C_4)]$

Examples- Training

Example 10: Reliability

(1) $\mathbb{P}(S) = \mathbb{P}(C_1 \cup C_2 \cup C_3)$
$= \mathbb{P}(C_1) + \mathbb{P}(C_2) + \mathbb{P}(C_3) - \mathbb{P}(C_1 \cap C_2) - \mathbb{P}(C_1 \cap C_3) - \mathbb{P}(C_2 \cap C_3) + \mathbb{P}(C_1 \cap C_2 \cap C_3)$
$= P_1 + P_2 + P_3 - P_1P_2 - P_1P_3 - P_2P_3 + P_1P_2P_3 = 1 - [(1 - P_1).(1 - P_2).(1 - P_3)]$
$= 1 - \mathbb{P}(C_1^c \cap C_2^c \cap C_3^c) \to S = (C_1^c \cap C_2^c \cap C_3^c)^c$

(2) $\mathbb{P}(S) = \mathbb{P}(C_1 \cap C_2 \cap C_3) = P_1.P_2.P_3$

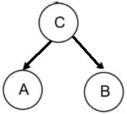
(3) $\mathbb{P}(S) = \mathbb{P}(C_3 \cap (C_1 \cup C_2)) = \mathbb{P}((C_3 \cap C_1) \cup (C_3 \cap C_2))$
$= \mathbb{P}(C_3 \cap C_1) + \mathbb{P}(C_3 \cap C_2) - \mathbb{P}(C_3 \cap C_1 \cap C_2) = P_3P_1 + P_3P_2 - P_1P_2P_3$
$P_3(P_1 + P_2 - P_1P_2) = P_3.[1 - (1 - P_1)(1 - P_2)]$

(4) $\mathbb{P}(S) = \mathbb{P}((C_1 \cap C_2) \cup C_3) = \mathbb{P}(C_1 \cap C_2) + \mathbb{P}(C_3) - \mathbb{P}(C_1 \cap C_2 \cap C_3) = P_1P_2 + P_3 - P_1P_2P_3$
$P_1P_2(1 - P_3) + P_3 \to S = (C_1 \cap C_2 \cap C_3^c) \cup C_3$

(5) $\mathbb{P}(S) = \mathbb{P}(C_5 \cap [(C_1 \cap C_2) \cup (C_3 \cap C_4)]) = \mathbb{P}(S) = \mathbb{P}[(C_5 \cap C_1 \cap C_2) \cup (C_5 \cap C_3 \cap C_4)]$
$= \mathbb{P}(C_5 \cap C_1 \cap C_2) + \mathbb{P}(C_5 \cap C_3 \cap C_4) - \mathbb{P}(C_5 \cap C_1 \cap C_2 \cap C_3 \cap C_4)$
$= P_5P_1P_2 + P_5P_3P_4 - P_5P_1P_2P_3P_4 = P_5[P_1P_2 + P_3P_4 - P_1P_2P_3P_4]$
$= P_5[1 - (1 - P_1P_2)(1 - P_3P_4)]$

Examples-Training

▷ Bayesian network / conditional independence from graphs



- We know: $\mathbb{P}(C)$, $\mathbb{P}(A|C)$, $\mathbb{P}(B|C)$ and A and B are conditionally independent

$$\begin{aligned}\mathbb{P}(A\cap B\cap C) &= \mathbb{P}(A\cap B|C).\mathbb{P}(C) \rightarrow \text{Bayes rule}\\ &= \underbrace{\mathbb{P}(A|C).\mathbb{P}(B|C)}_{\text{conditional independence}}.\mathbb{P}(C)\end{aligned}$$

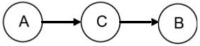
$$\boxed{\mathrm{P}(\mathrm{A}\cap B\cap C) = \mathbb{P}(A|C).\mathbb{P}(B|C).\mathbb{P}(C)}$$

- $$\begin{aligned}\mathbb{P}(A\cap B) &= \mathbb{P}(A\cap B\cap C) + \mathbb{P}(A\cap B\cap C^c) \rightarrow \text{total probability}\\ &= \mathbb{P}(A|C).\mathbb{P}(B|C).\mathbb{P}(C) + \mathbb{P}(A|C^c).\mathbb{P}(B|C^c).\mathbb{P}(C^c)\\ &\neq \mathbb{P}(A).\mathbb{P}(B) \rightarrow A \text{ and } B \text{ are not independent}\end{aligned}$$

$$(A \perp\!\!\!\perp B)|C \quad \text{But} \quad A \not\!\perp\!\!\!\perp B$$

Examples- Training

▷ Bayesian network / conditional independence from graphs



- We know: $\mathbb{P}(A)$, $\mathbb{P}(C|A)$, $\mathbb{P}(B|C)$ and A and B are conditionally independent

$$\begin{aligned}\mathbb{P}(A \cap B \cap C) &= \mathbb{P}(B|A \cap C).\mathbb{P}(A \cap C) \ \rightarrow \text{Bayes rule} \\ &= \underbrace{\mathbb{P}(B|C)}_{\text{conditional independence}} . \underbrace{\mathbb{P}(C|A).\mathbb{P}(A)}_{\text{Bayes rule}}\end{aligned}$$

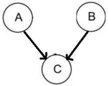
$$\boxed{\mathrm{P}(\mathrm{A} \cap B \cap C) = \mathbb{P}(A).\mathbb{P}(C|A).\mathbb{P}(B|C)}$$

- $$\begin{aligned}\mathbb{P}(A \cap B) &= \mathbb{P}(A \cap B \cap C) + \mathbb{P}(A \cap B \cap C^c) \ \rightarrow \text{total probability} \\ &= \mathbb{P}(A).\mathbb{P}(C|A).\mathbb{P}(B|C) + \mathbb{P}(A).\mathbb{P}(C^c|A).\mathbb{P}(B|C^c) \\ &\neq \mathbb{P}(A).\mathbb{P}(B) \rightarrow A \text{ and } B \text{ are not independent}\end{aligned}$$

$$(A \perp\!\!\!\perp B)|C \quad \text{But} \quad A \not\perp\!\!\!\perp B$$

Examples-Training

▷ Bayesian network / conditional independence from graphs



- We know: $\mathbb{P}(A)$, $\mathbb{P}(B)$, $\mathbb{P}(C|(A \cap B))$ and A and B are independent

$$\begin{aligned}\mathbb{P}(A \cap B \cap C) &= \mathbb{P}(C|A \cap B).\mathbb{P}(A \cap B) \quad \rightarrow \text{Bayes rule} \\ &= \mathbb{P}(C|(A \cap B)).\mathbb{P}(A).\mathbb{P}(B)\end{aligned}$$

$$\boxed{\mathrm{P}(\mathrm{A} \cap B \cap C) = \mathbb{P}(C|(A \cap B)).\mathbb{P}(A).\mathbb{P}(B)}$$

- $$\begin{aligned}\mathbb{P}((A \cap B)|C) &= \frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(C)} \quad \rightarrow \text{Bayes rule} \\ &= \frac{\mathbb{P}(C|(A \cap B)).\mathbb{P}(A).\mathbb{P}(B)}{\mathbb{P}(C)} \\ &\neq \frac{\mathbb{P}(A \cap C)}{\mathbb{P}(C)} + \frac{\mathbb{P}(B \cap C)}{\mathbb{P}(C)} = \mathbb{P}(A|C).\mathbb{P}(B|C)\end{aligned}$$

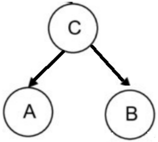
$$A \perp\!\!\!\perp B \quad \text{But} \quad (A \not\perp\!\!\!\perp B)|C$$

Examples- Training

Example 11: Suppose the event A and the event B are conditionally independent knowing C, and they are conditionally independent knowing C^c such that $\mathbb{P}(C) = 0.7$, $\mathbb{P}(A|C) = 0.4$, $\mathbb{P}(B|C) = 0.6$, $\mathbb{P}(A|C^c) = 0.3$ and $\mathbb{P}(B|C^c) = 0.2$. Show whether or not the pair {A,B} is independent.

Solution:

▷ We can construct the following table according to the information provided to us



	A	B	
C	$\mathbb{P}(A \cap C) = \mathbb{P}(A|C).\mathbb{P}(C) = 0.28$	$\mathbb{P}(B \cap C) = \mathbb{P}(B|C).\mathbb{P}(C) = 0.42$	0.7
C^c	$\mathbb{P}(A \cap C^c) = \mathbb{P}(A|C^c).\mathbb{P}(C^c) = 0.09$	$\mathbb{P}(B \cap C^c) = \mathbb{P}(B|C^c).\mathbb{P}(C^c) = 0.06$	0.3
	0.37	0.48	

- Note that, in this case A and B do not complete each other $\Rightarrow \mathbb{P}(A) + \mathbb{P}(B) \neq 1$

Examples-Training

$$\mathbb{P}(A) = \mathbb{P}(A \cap C) + \mathbb{P}(A \cap C^c) = 0.37$$

$$\mathbb{P}(B) = \mathbb{P}(B \cap C) + \mathbb{P}(B \cap C^c) = 0.48$$

$$\begin{aligned}
\mathbb{P}(A \cap B) &= \mathbb{P}(A \cap B \cap C) + \mathbb{P}(A \cap B \cap C^c) \\
&= \mathbb{P}(A \cap B|C).\mathbb{P}(C) + \mathbb{P}(A \cap B|C^c).\mathbb{P}(C^c) \\
&= \underbrace{\mathbb{P}(A|C).\mathbb{P}(B|C)}_{\text{conditional independence}}.\mathbb{P}(C) + \underbrace{\mathbb{P}(A|C^c).\mathbb{P}(B|C^c)}_{\text{conditional independence}}.\mathbb{P}(C^c) \\
&= (0.4).(0.6).(0.7) + (0.3).(0.2).(1 - 0.7) = 0.186
\end{aligned}$$

$\mathbb{P}(A \cap B) = 0.186 \neq \mathbb{P}(A).\mathbb{P}(B) = 0.177 \rightarrow$ A and B are dependent

Examples-Training

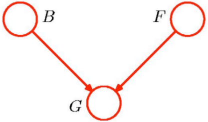
Example 12: (Bayesian network with conditional probability) Let the three binary variables

– Battery B $\rightarrow$ Charged (B=1) or Dead (B=0)

– Fuel Tank F $\rightarrow$ Full (F=1) or Empty (F=0)

– Guage Electric Fuel G $\rightarrow$ Indicates Full (G=1) or Empty (G=0)

– B and F are independent with prior probabilities (i.e; $\mathbb{P}(B \cap F) = \mathbb{P}(B).\mathbb{P}(F)$)



- We are given prior probabilities and one set of conditional probabilities

Battery
Prior Probabilities
$p(B)$

B	$p(B)$
1	0.9
0	0.1

Fuel
Prior Probabilities
$p(F)$

F	$p(F)$
1	0.9
0	0.1

Conditional probabilities of Guage
$p(G|B \cap F)$

B	F	$p(G=1)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

Examples- Training

1) Compute the probability that the guage reads full (i.e; $\mathbb{P}(G = 1)$) and deduce the probability that the guage reads empty (i.e; $\mathbb{P}(G = 0)$).
2) If the guage reads empty (G=0), what is the probability that the fuel tank being empty (i.e; $\mathbb{P}(F = 0|G = 0) = ?$).
2) If the guage reads empty (G=0), what is the probability that the battery being empty (i.e; $\mathbb{P}(B = 0|G = 0) = ?$).
4) Observing both fuel guage and battery. Suppose that the guage reads empty (G=0) and Battery is dead (B=0). What is the probability that Fuel tank is empty (i.e; $\mathbb{P}(F = 0|G = 0 \cap B = 0) = ?$).

Solution:

- $B = \{0, 1\} = \{(B = 0) \cup (B = 1)\} \Rightarrow \mathbb{P}(B) = \mathbb{P}(B = 0) + \mathbb{P}(B = 1) = 0.1 + 0.9 = 1$
- $F = \{0, 1\} = \{(F = 0) \cup (F = 1)\} \Rightarrow \mathbb{P}(F) = \mathbb{P}(F = 0) + \mathbb{P}(F = 1) = 0.1 + 0.9 = 1$
- $G = \{0, 1\} = \{(G = 0) \cup (G = 1)\} \Rightarrow \mathbb{P}(G) = \mathbb{P}(G = 0) + \mathbb{P}(G = 1) = 1$

Examples- Training

	$B=1\cap F=1$	$B=1\cap F=0$	$B=0\cap F=1$	$B=0\cap F=0$	Total
G=1	0.648	0.018	0.018	0.001	0.685
G=0	0.162	0.072	0.072	0.009	0.315
Total	0.81	0.09	0.09	0.01	1

$$\mathbb{P}(G=1\cap B=1\cap F=1)=\mathbb{P}(G=1|B=1\cap F=1).\mathbb{P}(B=1).\mathbb{P}(F=1)=0.64$$

$$\mathbb{P}(G=1\cap B=1\cap F=0)=\mathbb{P}(G=1|B=1\cap F=0).\mathbb{P}(B=1).\mathbb{P}(F=0)=0.018$$

$$\mathbb{P}(G=1\cap B=0\cap F=1)=\mathbb{P}(G=1|B=0\cap F=1).\mathbb{P}(B=0).\mathbb{P}(F=1)=0.018$$

$$\mathbb{P}(G=1\cap B=0\cap F=0)=\mathbb{P}(G=1|B=0\cap F=0).\mathbb{P}(B=0).\mathbb{P}(F=0)=0.001$$

$$\mathbb{P}(G=0\cap B=1\cap F=1)=[1-\mathbb{P}(G=1|B=1\cap F=1)].\mathbb{P}(B=1).\mathbb{P}(F=1)=0.162$$

$$\mathbb{P}(G=0\cap B=1\cap F=0)=[1-\mathbb{P}(G=1|B=1\cap F=0)].\mathbb{P}(B=1).\mathbb{P}(F=0)=0.072$$

$$\mathbb{P}(G=0\cap B=0\cap F=1)=[1-\mathbb{P}(G=0|B=0\cap F=1)].\mathbb{P}(B=0).\mathbb{P}(F=1)=0.072$$

$$\mathbb{P}(G=0\cap B=0\cap F=0)=[1-\mathbb{P}(G=1|B=0\cap F=0)].\mathbb{P}(B=0).\mathbb{P}(F=0)=0.009$$

Examples-Training

Battery
Prior Probabilities
p(B)

B	p(B)
1	0.9
0	0.1

Fuel
Prior Probabilities
p(F)

F	p(F)
1	0.9
0	0.1

Conditional probabilities of Guage
$p(G|B \cap F)$

B	F	p(G=1)
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

$$\mathbb{P}(G=1) = \underbrace{\mathbb{P}(G=1 \cap B=1 \cap F=1)}_{\mathbb{P}(G=1|B=1\cap F=1).\mathbb{P}(B=1\cap F=1)} + \underbrace{\mathbb{P}(G=1 \cap B=1 \cap F=0)}_{\mathbb{P}(G=1|B=1\cap F=0).\mathbb{P}(B=1\cap F=0)}$$

$$+ \underbrace{\mathbb{P}(G=1 \cap B=0 \cap F=1)}_{\mathbb{P}(G=1|B=0\cap F=1).\mathbb{P}(B=0\cap F=1)} + \underbrace{\mathbb{P}(G=1 \cap B=0 \cap F=0)}_{\mathbb{P}(G=1|B=0\cap F=0).\mathbb{P}(B=0\cap F=0)}$$

$$= \sum_{B=0,1} \sum_{F=0,1} \mathbb{P}(G=1|(B \cap F)).\mathbb{P}(B).\mathbb{P}(F)$$

$$\Rightarrow \mathbb{P}(G=1) = (0.8).(0.9)^2 + 2(0.2).(0.1) \times (0.9) + (0.1).(0.1)^2 = 0.685$$

$$\Rightarrow \mathbb{P}(G=0) = 1 - \mathbb{P}(G=1) = 0.315$$

Examples- Training

2) $$\mathbb{P}(F=0|G=0)=\frac{\mathbb{P}(G=0\cap F=0)}{\mathbb{P}(G=0)}$$

$$\mathbb{P}(G=0\ \cap\ F=0)=\underbrace{\mathbb{P}(G=0\ \cap\ F=0\ \cap B=0)}_{\mathbb{P}(G=0|(F=0\cap B=0)).\mathbb{P}(F=0).\mathbb{P}(B=0)}+\underbrace{\mathbb{P}(G=0\ \cap F=0\ \cap B=1)}_{\mathbb{P}(G=0|(F=0\cap B=1)).\mathbb{P}(F=0).\mathbb{P}(B=1)}$$

$$=(1-0.1).(0.1)^2+(1-0.2).(0.1).(0.9)=0.081$$

$$\Rightarrow \mathbb{P}(F=0|G=0)=\frac{\mathbb{P}(G=0\cap F=0)}{\mathbb{P}(G=0)}=\frac{0.081}{0.315}=0.257$$

- Note that, $\mathbb{P}(F=0|G=0)=0.257>\mathbb{P}(F=0)=0.1$

3) $$\mathbb{P}(B=0|G=0)=\frac{\mathbb{P}(G=0\cap B=0)}{\mathbb{P}(G=0)}$$

$$\mathbb{P}(G=0\ \cap\ B=0)=\underbrace{\mathbb{P}(G=0\ \cap\ B=0\ \cap F=0)}_{\mathbb{P}(G=0|(F=0\cap B=0)).\mathbb{P}(F=0).\mathbb{P}(B=0)}+\underbrace{\mathbb{P}(G=0\ \cap B=0\ \cap F=1)}_{\mathbb{P}(G=0|(F=1\cap B=0)).\mathbb{P}(F=1).\mathbb{P}(B=0)}$$

$$=(1-0.1).(0.1)^2+(1-0.2).(0.1).(0.9)=0.081$$

$$\Rightarrow \mathbb{P}(B=0|G=0)=\frac{\mathbb{P}(G=0\cap B=0)}{\mathbb{P}(G=0)}=\frac{0.081}{0.315}=0.257$$

Examples-Training

4) $\mathbb{P}(F|G\cap B) = \dfrac{\mathbb{P}(F\cap G\cap B)}{\mathbb{P}(G\cap B)} = \dfrac{\mathbb{P}(G|F\cap B).\mathbb{P}(F\cap B)}{\mathbb{P}(G|B).\mathbb{P}(B)} = \dfrac{\mathbb{P}(G|F\cap B).\mathbb{P}(F).\mathbb{P}(B)}{\mathbb{P}(G|B).\mathbb{P}(B)}$

$$\mathbb{P}(F=0|G=0\cap B=0) = \frac{\mathbb{P}(F=0\cap G=0\cap B=0)}{\mathbb{P}(G=0\cap B=0)}$$

$$= \frac{\mathbb{P}(G=0|F=0\cap B=0).\mathbb{P}(F=0\cap B=0)}{\mathbb{P}(G=0\cap B=0\cap F=0)+\mathbb{P}(G=0\cap B=0\cap F=1)} =$$

$$\frac{\mathbb{P}(G=0|F=0\cap B=0).\mathbb{P}(F=0).\mathbb{P}(B=0)}{[\mathbb{P}(G=0|B=0\cap F=0).\mathbb{P}(F=0)+\mathbb{P}(G=0|B=0\cap F=1).\mathbb{P}(F=1)].\mathbb{P}(B=0)}$$

$$= \frac{(1-0.1).(0.1).(0.1)}{[(1-0.1).(0.1).(0.1)]+[(1-0.2).(0.9).(0.1)]} = \frac{0.009}{0.009+0.072} = 0.1111$$

- Probability has decreased from 0.257 to 0.1111
- We can remark that

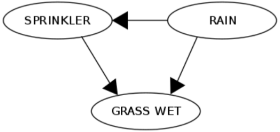
$$\mathbb{P}(B\cap F|G)|_0 = \frac{\mathbb{P}(G=0|B=0\cap F=0).\mathbb{P}(B=0).\mathbb{P}(F=0)}{\mathbb{P}(G=0)} = \frac{(0.9).(0.1)^2}{0.315} = 0.0286$$

$\neq \mathbb{P}(B=0|G=0).\mathbb{P}(F=0|G=0) = (0.257)^2 = 0.066 \rightarrow$ F and B are independent but they are not conditionally independent

Examples-Training:

Example 13: Two events can cause grass to be wet: an active sprinkler or rain. Rain has a direct effect on the use of the sprinkler (namely that when it rains, the sprinkler usually is not active). This situation can be modeled with a Bayesian network. Each variable has two possible values, T (for true) and F (for false).
Let G = "Grass wet (true/false)", S = "Sprinkler turned on (true/false)", and R = "Raining (true/false)".

RAIN	SPRINKLER T	SPRINKLER F
F	0.4	0.6
T	0.01	0.99



RAIN T	RAIN F
0.2	0.8

SPRINKLER	RAIN	GRASS WET T	GRASS WET F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

- The tables represent : $\mathbb{P}(G|R \cap S)$, $\mathbb{P}(S|R)$ and $\mathbb{P}(R)$.

Examples-Training:

- $\mathbb{P}(G = T \cap R = T \cap S = T) = \mathbb{P}(G = T|S = T \cap R = T).\mathbb{P}(S = T \cap R = T)$

$= \mathbb{P}(G = T|S = T \cap R = T).\mathbb{P}(S = T|R = T).\mathbb{P}(R = T) = 0.99 \times 0.01 \times 0.2 = 0.002$

- Using the same way of computation, we can get the following table

	$R = T \cap S = T$	$R = T \cap S = F$	$R = F \cap S = F$	$R = F \cap S = F$	Total
G=T	0.002	0.158	0.288	0	0.449
G=F	≈ 0	0.0396	0.032	0.48	0.551
Total	0.002	0.1976	0.32	0.48	1

- What is the probability that it is raining, given the grass is wet? $\rightarrow \mathbb{P}(R = T|G = T)$

$$\mathbb{P}(R = T|G = T) = \frac{\mathbb{P}(R = T \cap G = T)}{\mathbb{P}(G = T)} =$$

$$\frac{\mathbb{P}(R = T \cap G = T \cap S = T) + \mathbb{P}(R = T \cap G = T \cap S = F)}{\sum_{x,y \in \{T,F\}} \mathbb{P}(G = T \cap S = x \cap R = y)} = \frac{0.002 + 0.1585}{0.449} = 0.35725$$

Examples-Training:

- We can collect the following measurements:

$\mathbb{P}(R = T \cap S = T) = \mathbb{P}(S = T|R = T).\mathbb{P}(R = T) = 0.01 \times 0.2 = 0.002$

- Using the same way of computation, we can get the following table

	$R = T$	$R = F$	Total
S=T	0.002	0.32	0.322
S=F	0.198	0.48	0.678
Total	0.2	0.8	1

- What is the probability that it is raining and the sprinkler turned on given the grass is wet? $\rightarrow \mathbb{P}(R = T \cap S = T|G = T)$

- $\mathbb{P}(R = T \cap S = T|G = T) = \dfrac{\mathbb{P}(R = T \cap S = T \cap G = T)}{\mathbb{P}(G = T)} = \dfrac{0.002}{0.449} = 0.0045$

$\neq \mathbb{P}(R = T \cap S = T) \rightarrow$ R and S are conditionally dependent.

- $\mathbb{P}(S = T \cap R = T) = 0.002 \neq \mathbb{P}(S = T).\mathbb{P}(R = T) = 0.322 \times 0.2 = 0.0644$

$\rightarrow$ R and S are dependent.

Examples-Training:

Example 14:

Show that if event A and event B are independent by knowing that the event C occurs (i.e; conditionally independent $\mathbb{P}(A \cap B|C) = \mathbb{P}(A|C).\mathbb{P}(B|C)$) we have the following relationships:

1) $\mathbb{P}(A|B \cap C) = \mathbb{P}(A|C)$
2) $\mathbb{P}(A|B^c \cap C) = \mathbb{P}(A|C)$
3) $\mathbb{P}(A^c|B \cap C) = \mathbb{P}(A^c|C)$
4) $\mathbb{P}(A^c|B^c \cap C) = \mathbb{P}(A^c|C)$
5) $\mathbb{P}(A^c \cap B^c|C) = \mathbb{P}(A^c|C).\mathbb{P}(B^c|C)$

Examples-Training:

Suppose that A and B are independent by knowing that the event C occurs
$\rightarrow \mathbb{P}(A \cap B|C) = \mathbb{P}(A|C).\mathbb{P}(B|C)$.

1) $\mathbb{P}(A|B \cap C) = \mathbb{P}(A|C)$

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

2) $\mathbb{P}(A|B^c \cap C) = \mathbb{P}(A|C)$

$$\mathbb{P}(A|B^c \cap C) = \frac{\mathbb{P}(A \cap B^c \cap C)}{\mathbb{P}(B^c \cap C)} = \frac{\mathbb{P}(A \cap B^c|C).\mathbb{P}(C)}{\mathbb{P}(B^c|C).\mathbb{P}(C)} = \frac{\mathbb{P}(A \cap B^c|C)}{\mathbb{P}(B^c|C)} = \frac{\mathbb{P}(A|C).\mathbb{P}(B^c|C)}{\mathbb{P}(B^c|C)}$$
$$= \mathbb{P}(A|C)$$

3) $\mathbb{P}(A^c|B \cap C) = \mathbb{P}(A^c|C)$

$$\mathbb{P}(A^c|B \cap C) = \frac{\mathbb{P}(A^c \cap B \cap C)}{\mathbb{P}(B \cap C)} = \frac{\mathbb{P}(A^c \cap B|C).\mathbb{P}(C)}{\mathbb{P}(B|C).\mathbb{P}(C)} = \frac{\mathbb{P}(A^c \cap B|C)}{\mathbb{P}(B|C)} = \frac{\mathbb{P}(A^c|C).\mathbb{P}(B|C)}{\mathbb{P}(B|C)}$$
$$= \mathbb{P}(A^c|C)$$

Examples-Training:

$$4)\ \mathbb{P}(A^c|B^c\cap C) = \frac{\mathbb{P}(A^c\cap B^c\cap C)}{\mathbb{P}(B^c\cap C)} = \frac{\mathbb{P}(A^c\cap B^c|C).\mathbb{P}(C)}{\mathbb{P}(B^c|C).\mathbb{P}(C)} = \frac{\mathbb{P}(A^c\cap B^c|C)}{\mathbb{P}(B^c|C)} = \frac{\mathbb{P}(A^c|C).\mathbb{P}(B^c|C)}{\mathbb{P}(B^c|C)}$$

$$= \mathbb{P}(A^c|C)$$

$$5)\ \mathbb{P}(A^c\cap B^c|C) = \mathbb{P}((A\cup B)^c|C) = 1-\mathbb{P}((A\cup B)^c|C)$$

$$= 1-\frac{\mathbb{P}((A\cup B)\cap C)}{\mathbb{P}(C)} = 1-\frac{\mathbb{P}((A\cap C)\cup(B\cap C))}{\mathbb{P}(C)}$$

$$= 1-\frac{\mathbb{P}(A\cap C)+\mathbb{P}(B\cap C)-\mathbb{P}(A\cap B\cap C)}{\mathbb{P}(C)}$$

$$= 1-\frac{\mathbb{P}(A\cap C)}{\mathbb{P}(C)}-\frac{\mathbb{P}(B\cap C)}{\mathbb{P}(C)}+\frac{\mathbb{P}(A\cap B\cap C)}{\mathbb{P}(C)}$$

$$= 1-\mathbb{P}(A|C)-\mathbb{P}(B|C)+\mathbb{P}(A\cap B|C)$$

$$= 1-\mathbb{P}(A|C)-\mathbb{P}(B|C)+\mathbb{P}(A|C).\mathbb{P}(B|C)$$

$$= [1-\mathbb{P}(A|C)].[1-\mathbb{P}(B|C)]$$

$$= \mathbb{P}(A^c|C).\mathbb{P}(B^c|C)$$

Axioms

Axioms-Conditional probability:

1) $\mathbb{P}(A|B) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \dfrac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)} \rightarrow$ Conditional probability and Baye's rule
2) $\mathbb{P}(A \cap B) = \mathbb{P}(A|B).\mathbb{P}(B)$ AND $\mathbb{P}(A \cap B) = \mathbb{P}(B|A).\mathbb{P}(A)$
3) A and B are **independent** if $\mathbb{P}(A|B) = \mathbb{P}(A)$ OR $\mathbb{P}(A \cap B) = \mathbb{P}(A).\mathbb{P}(B)$
4) $\mathbb{P}(A|B) = 1 - \mathbb{P}(A^c|B)$
5) $\mathbb{P}(A) = \mathbb{P}[(A \cap \square) \cup (A \cap \square^c)] = \mathbb{P}(A \cap \square) + \mathbb{P}(A \cap \square^c) = \mathbb{P}(A|\square).\mathbb{P}(\square) + \mathbb{P}(A|\square^c).\mathbb{P}(\square^c) \rightarrow$ The total probability $\rightarrow \square$ can be any set)
6) A and B are independent knowing that C occurs if

$$\mathbb{P}((A \cap B)|C) = \mathbb{P}(A|C).\mathbb{P}(B|C) \quad \rightarrow \quad \text{conditional independence}$$