CS70 @ UC Berkeley, Spring 2026

Lecture 18

Combinations of Events, Mutual Independence

March 31, 2026

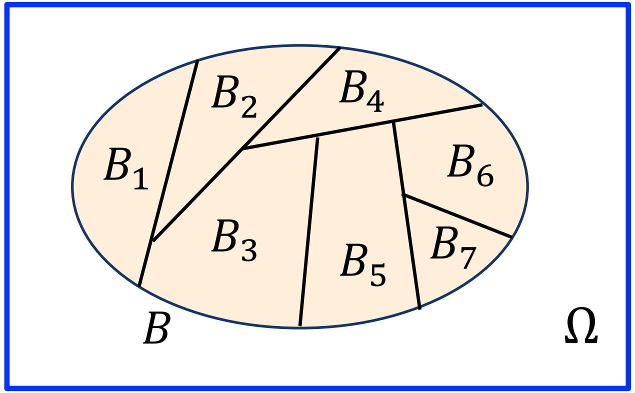
Last Lecture

Definition (Partition): An event $B \subseteq \Omega$ is said to be partitioned into n events $B_1, \dots, B_n$ if

1. $B = B_1 \cup \cdots \cup B_n$,

2. $B_i \cap B_j = \emptyset$, for all $i \neq j$ (that is, $B_1, \dots, B_n$ are mutually exclusive).

More generally, infinitely many mutually exclusive B_ks may be involved ($B = \cup_{k=1}^{\infty} B_k$).



1. (**Non-negativity**) $\mathbb{P}(A) \geq 0$, for all $A \subseteq \Omega$.
2. (**Countable Additivity**) If $B_1, B_2, B_3, \dots$ is a partition of B, then

$$\mathbb{P}(B) = \sum_{k=1}^{\infty} \mathbb{P}(B_k)$$

3. (**Normalization**) $\mathbb{P}(\Omega) = 1$.

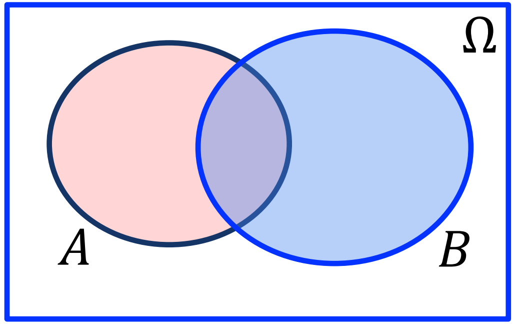
How about $\mathbb{P}(\emptyset)$?

Last Lecture

- Consider two events $A, B \subseteq \Omega$.
- Conditioning on event B changes the probability space from $(\Omega, \mathbb{P})$ to $(B, \mathbb{P}_B)$.
- Conditional probability of A given B:

$$\mathbb{P}(A|B) \coloneqq \mathbb{P}_B(A \cap B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \leq 1.$$

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



Theorem (Law of Total Probability or Total Probability Rule). Suppose $B_1, \ldots, B_n$ is a partition of Ω. Then, for any $A \subseteq \Omega$,

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i)$$

a.k.a. rule of average conditional probabilities.

Which Coin? (Last Lecture)

	Coin 1	**Coin 2**	**Coin 3**
Probability of H	p_1	p_2	p_3
Probability of T	$1-p_1$	$1-p_2$	$1-p_3$

- Pick a coin uniformly at random (u.a.r.) and toss it once.
- **Q:** What is the probability of observing Heads?
- $\Omega = \{(1,H),(1,T),(2,H),(2,T),(3,H),(3,T)\}$
- Event C_i = "coin i is chosen" $= \{(i,H),(i,T)\}$
- Note: $C_1 \cup C_2 \cup C_3 = \Omega$ and $C_i \cap C_j = \emptyset$ for $i \neq j$, so C_1, C_2, C_3 is a partition of Ω.
- $\mathbb{P}(H) = \sum_{i=1}^{3} \mathbb{P}(H|C_i)\mathbb{P}(C_i) = \frac{1}{3}(p_1+p_2+p_3)$

$$\uparrow$$

Prior (belief before any observation)

Pick a coin u.a.r. and toss it twice.

Is $\mathbb{P}(HH) = \mathbb{P}(H)\mathbb{P}(H)$?

$$\mathbb{P}(HH) = \sum_{i=1}^{3} \mathbb{P}(HH|C_i)\mathbb{P}(C_i)$$

$$= \frac{1}{3}\sum_{i=1}^{3} p_i^2$$

$$\neq [\mathbb{P}(H)]^2$$

$$= \frac{1}{9}(p_1+p_2+p_3)^2$$

However,

$$\mathbb{P}(HH|C_i) = \mathbb{P}(H|C_i)\mathbb{P}(H|C_i)$$

Which Coin?

	Coin 1	**Coin 2**	**Coin 3**
Probability of H	p_1	p_2	p_3
Probability of T	$1 - p_1$	$1 - p_2$	$1 - p_3$

- Pick a coin uniformly at random (u.a.r.) and toss it once.
- **Q:** What is the probability of observing Heads?
- $\Omega = \{(1,H),(1,T),(2,H),(2,T),(3,H),(3,T)\}$
- Event C_i = "coin i is chosen" $= \{(i,H),(i,T)\}$
- Note: $C_1 \cup C_2 \cup C_3 = \Omega$ and $C_i \cap C_j = \emptyset$ for $i \neq j$, so C_1, C_2, C_3 is a partition of Ω.
- $\mathbb{P}(H) = \sum_{i=1}^{3} \mathbb{P}(H|C_i)\mathbb{P}(C_i) = \frac{1}{3}(p_1 + p_2 + p_3)$

Prior (belief before any observation)

Q: What is $\mathbb{P}[C_i|H]$?

Posterior (updated belief)

$$\mathbb{P}[C_i|H] = \frac{\mathbb{P}[C_i \cap H]}{\mathbb{P}(H)}$$

Bayes' Rule $$= \frac{\mathbb{P}(H|C_i)\mathbb{P}(C_i)}{\mathbb{P}(H)}$$

$\mathbb{P}(C_i) = \frac{1}{3}$ for all $i = 1,2,3$.

$$\mathbb{P}[C_i|H] = \frac{p_i}{p_1 + p_2 + p_3}$$

Which Coin?

	Coin 1	**Coin 2**	**Coin 3**
Probability of H	p_1	p_2	p_3
Probability of T	$1 - p_1$	$1 - p_2$	$1 - p_3$

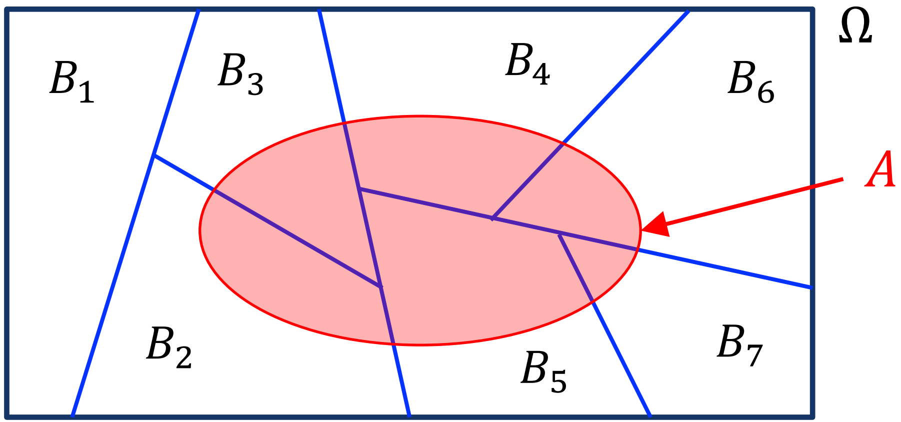
Posterior probabilities

$$\mathbb{P}[C_i|\text{Data}] = \frac{\mathbb{P}(\text{Data}|C_i)\mathbb{P}(C_i)}{\sum_{j=1}^{3}\mathbb{P}(\text{Data}|C_j)\mathbb{P}(C_j)}$$

For $p_1 = 0.4,\ p_2 = 0.5,\ p_3 = 0.6$:

Data	$\mathbb{P}[C_1|\text{Data}]$	$\mathbb{P}[C_2|\text{Data}]$	$\mathbb{P}[C_3|\text{Data}]$
H	0.27	0.33	0.40
HT	0.33	0.34	0.33
HTTTH	0.39	0.35	0.26
HTTTHTTHTT	0.57	0.31	0.11
HTTTHTTHTTTHTHT	0.62	0.30	0.08
HTTTHTTHTTTHTHTTTTTTTHHHHTHHTHTHHTTTHTTT	0.76	0.23	0.01
HTTTHTTHTTTHTHTTTTTTTHHHHTHHTHTHHTTTHTTT *HTTTTHTTTHHTTTHHHTHTTHTTHTTTHHTTTHTTTTTH*	0.94	0.06	0.00

Bayes' Rule



Let $B_1, \dots, B_n$ be a partition of Ω. Then for any $A \subseteq \Omega$ with $\mathbb{P}(A) > 0$,

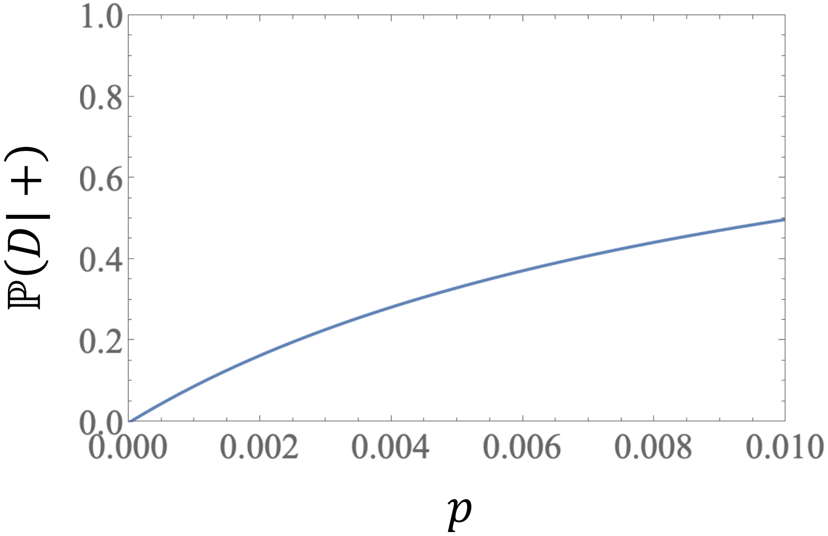
$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum_{j=1}^{n} \mathbb{P}\left(A|B_j\right)\mathbb{P}(B_j)}$$

Disease Screening

- Suppose Bob gets tested for a disease and the test result comes back “positive”
- Should Bob worry?
- Event D = has the disease
- Test result $\in \{+,-\}$
- $\mathbb{P}(D|+) = \frac{\mathbb{P}(+|D)\mathbb{P}(D)}{\mathbb{P}(+|D)\mathbb{P}(D)+\mathbb{P}(+|D^c)\mathbb{P}(D^c)}$
- $\mathbb{P}(D) = p$ (prior, disease prevalence)
- $\mathbb{P}(D^c) = 1 - p$
- $\mathbb{P}(+|D^c)$ = False Positive Rate
- $\mathbb{P}(-|D)$ = False Negative Rate

$$\boxed{\mathbb{P}(D|+) = \frac{(1-\mathrm{FNR})p}{(1-\mathrm{FNR})p + \mathrm{FPR}(1-p)}}$$

For $\mathbb{P}(+|D^c) = 0.01$ and $\mathbb{P}(-|D) = 0.01$:

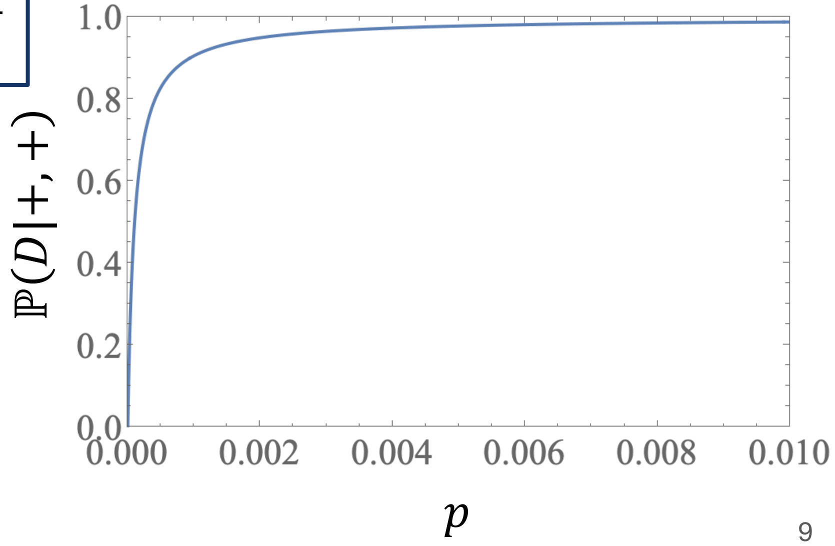


Disease Screening (two tests)

- Suppose Bob gets tested again and it also returns “positive”.
- Should Bob worry?
- Assuming that the two test results are conditionally independent given the disease status, and that they have the same FNR and FPR, we obtain

$$\boxed{\mathbb{P}(D|+,+) = \frac{(1-\text{FNR})^2 p}{(1-\text{FNR})^2 p + \text{FPR}^2(1-p)}}$$

For $\mathbb{P}(+|D^c) = 0.01$ and $\mathbb{P}(-|D) = 0.01$:



Recall Independence from Lecture 17

Consider any two events $A, B \subseteq \Omega$ and suppose the chance of A does not depend on whether or not B occurs, i.e., $\mathbb{P}(A|B) = \mathbb{P}(A|B^c)$. Then,

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

Note: we will use B^c and $\overline{B}$ interchangeably to denote "Not B"

When this condition holds, we say that A and B are independent events.

$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ implies all of the following:

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

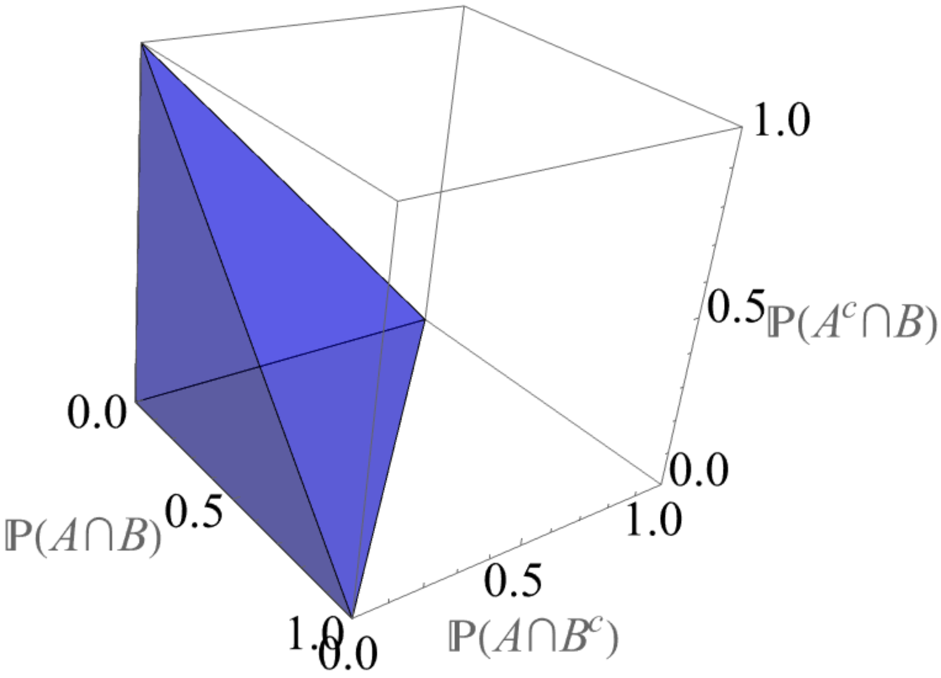
Proof of 1:

$\mathbb{P}(A^c \cap B) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$ Since $A^c \cap B$ and $A \cap B$ partition B, and by additivity.

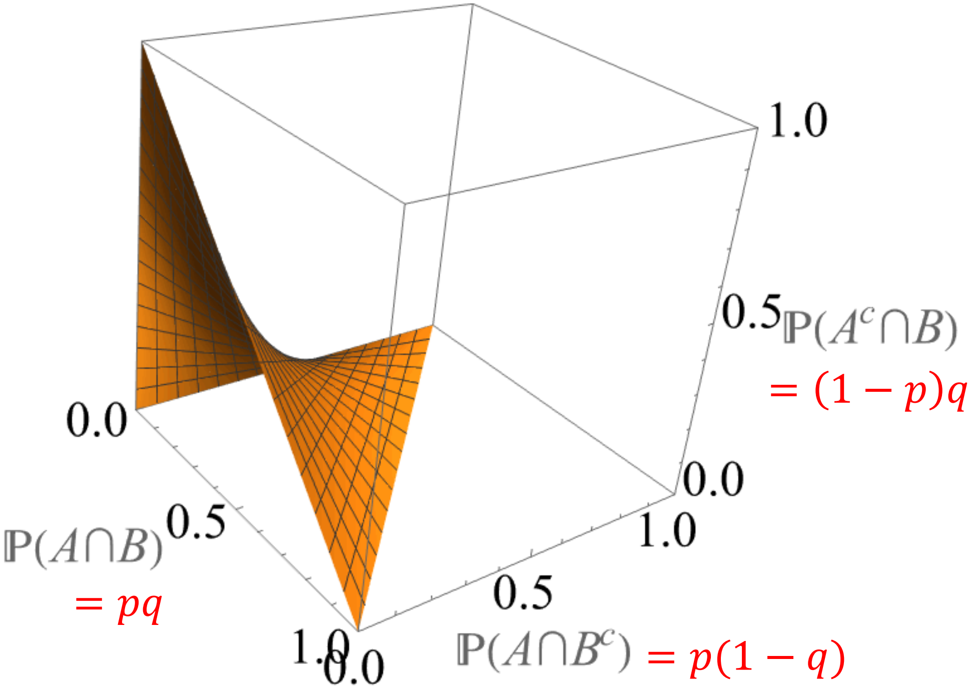
$= \mathbb{P}(B) - \mathbb{P}(A)\mathbb{P}(B)$ By assumption.

$= \mathbb{P}(B)\,\mathbb{P}(A^c)$ Since A and A^c partition Ω, and $\mathbb{P}(A) + \mathbb{P}(A^c) = \mathbb{P}(\Omega) = 1$ ☐

Space of Probability Measures for Two Events



Every point on or inside this 3-dimensional tetrahedron corresponds to a valid probability measure on A and B.



$= (1 - p)q$

$= pq$

$= p(1 - q)$

Every point on this 2-dimensional surface corresponds to a valid independent probability measure on A and B.

Mutual Independence for n Events

Does $\mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3)$ imply $\mathbb{P}(A_1^c \cap A_2^c \cap A_3) = \mathbb{P}(A_1^c)\mathbb{P}(A_2^c)\mathbb{P}(A_3)$ and so on? **No!**

Definition (Independence of n events). Events $A_1, \ldots, A_n$ are mutually independent if for all subsets $I \subseteq \{1, \ldots, n\}$ with $|I| \geq 2$,

$$\mathbb{P}\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \mathbb{P}(A_i).$$

These define $2^n - n - 1$ equations.

For $n = 3, A_1, A_2, A_3$ are mutually independent if

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

The dimension of valid probability measures on n events is $2^n - 1$.

The dimension of independent probability measures on n events is $2^n - 1 - (2^n - n - 1) = n$

Mutual Independence for n Events

Definition (Independence of n events). Equivalently, **events** $A_1, \dots, A_n$ are mutually independent if for all $B_i \in \{A_i, A_i^c\}$, $i = 1, \dots, n$,

$$\mathbb{P}(B_1 \cap \cdots \cap B_n) = \prod_{i=1}^{n} \mathbb{P}(B_i).$$

These define 2^n equations, but $(n + 1)$ of them are redundant.

For $n = 3$, suppose

$$\begin{aligned}
\mathbb{P}(A_1 \cap A_2 \cap A_3) &= \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3) \\
\mathbb{P}(A_1^c \cap A_2 \cap A_3) &= \mathbb{P}(A_1^c)\mathbb{P}(A_2)\mathbb{P}(A_3) \\
\mathbb{P}(A_1 \cap A_2^c \cap A_3) &= \mathbb{P}(A_1)\mathbb{P}(A_2^c)\mathbb{P}(A_3) \\
\mathbb{P}(A_1 \cap A_2 \cap A_3^c) &= \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3^c)
\end{aligned}$$

Then, the above equations imply

$$\begin{aligned}
\mathbb{P}(A_1^c \cap A_2^c \cap A_3^c) &= \mathbb{P}(A_1^c)\mathbb{P}(A_2^c)\mathbb{P}(A_3^c) \\
\mathbb{P}(A_1 \cap A_2^c \cap A_3^c) &= \mathbb{P}(A_1)\mathbb{P}(A_2^c)\mathbb{P}(A_3^c) \\
\mathbb{P}(A_1^c \cap A_2 \cap A_3^c) &= \mathbb{P}(A_1^c)\mathbb{P}(A_2)\mathbb{P}(A_3^c) \\
\mathbb{P}(A_1^c \cap A_2^c \cap A_3) &= \mathbb{P}(A_1^c)\mathbb{P}(A_2^c)\mathbb{P}(A_3)
\end{aligned}$$

Independence for n Events

Suppose

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

$$\mathbb{P}(A_1 \cap A_2 \cap A_3) + \mathbb{P}(A_1 \cap A_2^c \cap A_3) + \mathbb{P}(A_1 \cap A_2 \cap A_3^c) + \mathbb{P}(A_1 \cap A_2^c \cap A_3^c) = \mathbb{P}(A_1)$$
$$\Longrightarrow \mathbb{P}(A_1 \cap A_2^c \cap A_3^c) = \mathbb{P}(A_1)\mathbb{P}(A_2^c)\mathbb{P}(A_3^c)$$

$$\mathbb{P}(A_1 \cap A_2 \cap A_3) + \mathbb{P}(A_1^c \cap A_2 \cap A_3) + \mathbb{P}(A_1 \cap A_2 \cap A_3^c) + \mathbb{P}(A_1^c \cap A_2 \cap A_3^c) = \mathbb{P}(A_2)$$
$$\Longrightarrow \mathbb{P}(A_1^c \cap A_2 \cap A_3^c) = \mathbb{P}(A_1^c)\mathbb{P}(A_2)\mathbb{P}(A_3^c)$$

$$\mathbb{P}(A_1 \cap A_2 \cap A_3) + \mathbb{P}(A_1^c \cap A_2 \cap A_3) + \mathbb{P}(A_1 \cap A_2^c \cap A_3) + \mathbb{P}(A_1^c \cap A_2^c \cap A_3) = \mathbb{P}(A_3)$$
$$\Longrightarrow \mathbb{P}(A_1^c \cap A_2^c \cap A_3) = \mathbb{P}(A_1^c)\mathbb{P}(A_2^c)\mathbb{P}(A_3)$$

$$\mathbb{P}(A_1 \cap A_2 \cap A_3) + \mathbb{P}(A_1^c \cap A_2 \cap A_3) + \mathbb{P}(A_1 \cap A_2^c \cap A_3) + \mathbb{P}(A_1 \cap A_2 \cap A_3^c) +$$
$$\mathbb{P}(A_1 \cap A_2^c \cap A_3^c) + \mathbb{P}(A_1^c \cap A_2 \cap A_3^c) + \mathbb{P}(A_1^c \cap A_2^c \cap A_3) + \mathbb{P}(A_1^c \cap A_2^c \cap A_3^c) = 1$$
$$\Longrightarrow \mathbb{P}(A_1^c \cap A_2^c \cap A_3^c) = \mathbb{P}(A_1^c)\mathbb{P}(A_2^c)\mathbb{P}(A_3^c)$$

Pairwise Independence vs. Mutual Independence

Definition (Pairwise Independence). Events $A_1, \ldots, A_n$ are said to be pairwise independent if

$$\mathbb{P}\left(A_i \cap A_j\right) = \mathbb{P}(A_i)\mathbb{P}(A_j)$$

for all i, j such that $i \neq j$.

Does pairwise independence imply mutual independence? **NO!**

- Toss a fair coin three times.
- $\Omega = \{(H,H,H),(H,H,T),(H,T,H),\ (T,H,H),(H,T,T),(T,H,T),(T,T,H),(T,T,T)\}$
- Let E_{ij} denote the event that the ith and the jth tosses show the same face.
- Then,

$$\mathbb{P}(E_{12}) = \mathbb{P}(E_{13}) = \mathbb{P}(E_{23}) = \frac{1}{2}$$

$$\mathbb{P}(E_{12} \cap E_{23}) = \frac{1}{4} = \mathbb{P}(E_{12})\mathbb{P}(E_{23})$$

$$\mathbb{P}(E_{12} \cap E_{13}) = \frac{1}{4} = \mathbb{P}(E_{12})\mathbb{P}(E_{13})$$

$$\mathbb{P}(E_{13} \cap E_{23}) = \frac{1}{4} = \mathbb{P}(E_{13})\mathbb{P}(E_{23})$$

$$E_{12} = \{(H,H,H),(H,H,T),(T,T,H),(T,T,T)\}$$
$$E_{13} = \{(H,H,H),(H,T,H),(T,H,T),(T,T,T)\}$$
$$E_{23} = \{(H,H,H),(T,H,H),(H,T,T),(T,T,T)\}$$

However,

$$\mathbb{P}(E_{12} \cap E_{23} \cap E_{13}) = \frac{1}{4} \neq \mathbb{P}(E_{12})\mathbb{P}(E_{23})\mathbb{P}(E_{13}) = \frac{1}{8}$$

Product Rule

Theorem (Product Rule). For any events $A_1, \ldots, A_n$ on the same probability space,
$\mathbb{P}(A_1 \cap \cdots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2) \cdots \mathbb{P}(A_n|A_1 \cap \cdots \cap A_{n-1})$.

Proof. By induction on n.
Base case: For $n = 1$, $\mathbb{P}(A_1) = \mathbb{P}(A_1)$.
Induction hypothesis: Assume true for all $n \leq k$, where k is an arbitrary positive integer.
Then,

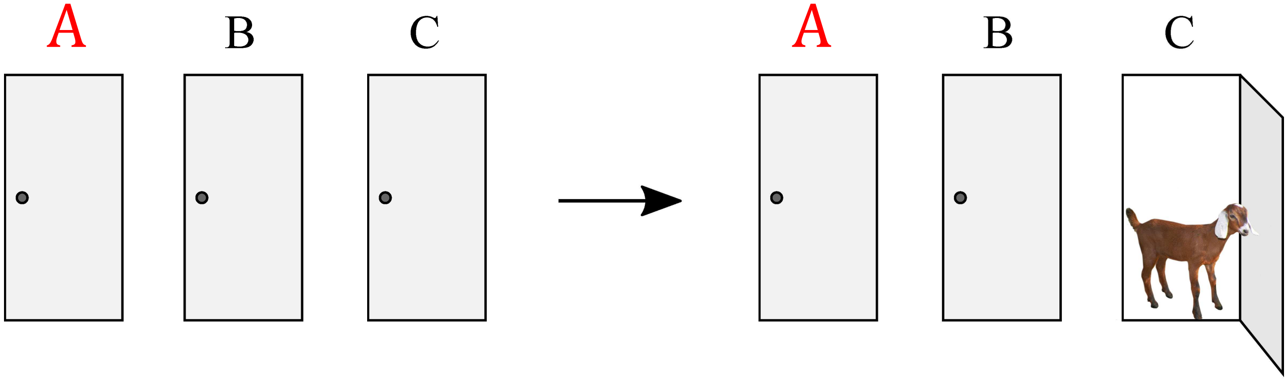
$$\mathbb{P}\left(\cap_{i=1}^{k+1} A_i\right) = \mathbb{P}\left(\left(\cap_{i=1}^{k} A_i\right) \cap A_{k+1}\right) = \mathbb{P}\left(\cap_{i=1}^{k} A_i\right) \mathbb{P}\left(A_{k+1} \middle| \cap_{i=1}^{k} A_i\right) \quad \text{by definition of conditional probability}$$

$$= \underbrace{\mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2) \cdots \mathbb{P}\left(A_k \middle| \cap_{i=1}^{k-1} A_i\right)}_{\text{by induction hypothesis}} \mathbb{P}\left(A_{k+1} \middle| \cap_{i=1}^{k} A_i\right)$$

$\square$

Monty Hall Revisited

- A famous game show from 1970s, hosted by Monty Hall.
- One door has a car behind it. The other two doors have goats.



Question: Does the contestant have a better chance of winning if he/she switches doors?

- The contestant picks a door (A in this example), but doesn't open it.
- Monty Hall's assistant (Carol) opens one of the two remaining doors to reveal a goat (since Carol knows where the prize is, she can always do this).
- The contestant is then given the option of sticking with his/her current door, or switching to the other unopened one.

Monty Hall Revisited

- P_i = the prize is behind door i
- C_i = the contestant chooses door i
- R_i = Carol reveals door i
- $\mathbb{P}(P_2 \cap C_1 \cap R_3) = \mathbb{P}(P_2)\mathbb{P}(C_1|P_2)\mathbb{P}(R_3|P_2 \cap C_1) = \frac{1}{3} \times \frac{1}{3} \times 1 = \frac{1}{9}.$
- $\mathbb{P}(P_1 \cap C_1 \cap R_3) = \mathbb{P}(P_1)\mathbb{P}(C_1|P_1)\mathbb{P}(R_3|P_1 \cap C_1) = \frac{1}{3} \times \frac{1}{3} \times \frac{1}{2} = \frac{1}{18}.$
- Since P_1, P_2, P_3 partition Ω,

$$\mathbb{P}(C_1 \cap R_3) = \underset{\frac{1}{18}}{\mathbb{P}(P_1 \cap C_1 \cap R_3)} + \underset{\frac{1}{9}}{\mathbb{P}(P_2 \cap C_1 \cap R_3)} + \underset{0}{\mathbb{P}(P_3 \cap C_1 \cap R_3)} = \frac{1}{6}$$

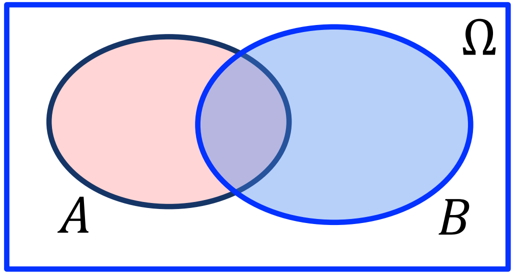
$$\mathbb{P}(P_2|C_1 \cap R_3) = \frac{\mathbb{P}(P_2 \cap C_1 \cap R_3)}{\mathbb{P}(C_1 \cap R_3)} = \frac{\frac{1}{9}}{\frac{1}{6}} = \frac{2}{3}$$

$$\mathbb{P}(P_1|C_1 \cap R_3) = \frac{\mathbb{P}(P_1 \cap C_1 \cap R_3)}{\mathbb{P}(C_1 \cap R_3)} = \frac{\frac{1}{18}}{\frac{1}{6}} = \frac{1}{3}$$

Hence, the contestant would be better off switching to door 3.

Union of Events

Let A, B be two events on the same probability space.



$$\mathbb{P}(A) = \mathbb{P}(A\backslash(A \cap B)) + \mathbb{P}(A \cap B)$$

$$\mathbb{P}(B) = \mathbb{P}(B\backslash(A \cap B)) + \mathbb{P}(A \cap B)$$

$$\begin{aligned}\mathbb{P}(A \cup B) &= \mathbb{P}(A\backslash(A \cap B)) + \mathbb{P}(B\backslash(A \cap B)) + \mathbb{P}(A \cap B) \\ &= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)\end{aligned}$$

Recall the Inclusion-Exclusion formula from Lecture 15:

Theorem 5 (Inclusion-Exclusion). *Let* $A_1, \ldots, A_n$ *be arbitrary subsets of the same set* Ω. *Then,*

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{k=1}^{n} (-1)^{k-1} \sum_{S \subseteq \{1,\ldots,n\} : |S| = k} \left| \bigcap_{i \in S} A_i \right|.$$

Replace $|E|$ with $\mathbb{P}(E)$ in both sides to obtain the Inclusion-Exclusion formula for $\mathbb{P}(A_1 \cup \cdots \cup A_n)$.

Union of Events

Remarks:

1. If $A_1, \dots, A_n$ are mutually exclusive ($A_i \cap A_j = \emptyset$ for all $i \neq j$) events, then
$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i).$$
By additivity

2. **Union Bound:** For all events $A_1, \dots, A_n$ on the same probability space,
$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} \mathbb{P}(A_i).$$

This inequality can be proved using induction.
This bound has many applications in Computer Science.