# GSOE9210 Engineering Decisions

Victor Jauregui

vicj@cse.unsw.edu.au
www.cse.unsw.edu.au/~gs9210



## Probability

# Outline

1. Measuring uncertainty
    - Probability





# The ontology of uncertainty

**Example (Tossing a coin)**

After a coin is tossed, the top-most face may be 'heads' or 'tails'.

- Tossing a coin is a *random process* or *chance experiment*
- Each toss of the coin corresponds to a *trial* of the experiment
- 'Heads' is one among the experiment's *possible individual outcomes*
- 'Tossing heads' is an observable *event*

## The ontology of uncertainty

**Example (Rolling a die)**

After a die is rolled, the top-most face may be any of a *one*, *two*, . . . , *six*.

- Observable events include rolling a six; a three; an even number; a number greater than two
- Its not always the case that individual outcomes can be observed; *e.g.*, experimental error





## The ontology of uncertainty

In a decision problem:

- Each state (of nature) and outcome/consequence in a decision problem corresponds to an event
- States and outcomes are assumed to be mutually exclusive
- Experimental data and observations can provide information about 'likelihoods' of certain events (states/outcomes)
- Likelihood estimates may be objective or subjective, and may change depending on the agent's epistemic state

## Frequency interpretation of probability

Outcomes:

$$\underbrace{t, h, h, t, t, h, t}_{n}, h, t, h, \ldots$$

**Definition (Frequency interpretation of probability)**

The *probability* of an event, $E$, in an experiment of chance, is the limit of the average occurrences of $E$ over any sequence of indefinitely many trials; *i.e.*,

$$P(E) = \lim_{n\to\infty} \frac{n_E}{n}$$

where $n_E$ is the number of occurrences of event $E$ in the first $n$ trials.

- *e.g.*, For event $H$: $\frac{0}{1}, \frac{1}{2}, \frac{2}{3}, \frac{2}{4}, \frac{2}{5}, \frac{3}{6}, \frac{3}{7}, \ldots, \frac{n_H}{n}, \ldots$
- What is $P(H)$ for this experiment?





## Bayesian interpretation

- The frequency interpretation assumes *repeatable* experiments/processes
- It is *objective*: a fixed property of the underlying random process/system; the same for every agent irrespective of the particular sequence of trials or of the epistemic state of the agents
- Problem: What is the ‘probability’ that Germany will win the 2018 football world cup?

**Definition (Bayesian probability)**

Probability is an agent’s *subjective degree of belief* in the occurrence of an event.

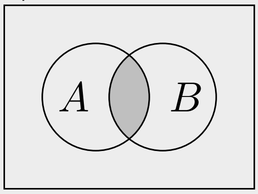
- An agent’s subjective probability depends on the agent’s epistemic state (knowledge and beliefs)

## Set operations

Assume $A, B \subseteq \Omega$.

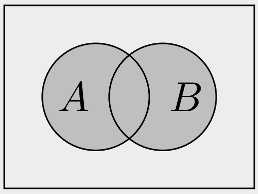
- The *intersection* of sets $A$ and $B$, written $A \cap B$ (or just $AB$), is the set of all elements common to both $A$ and $B$; *i.e.*,

  $$A \cap B = \{x \in \Omega \mid x \in A \;\; \& \;\; x \in B\}$$



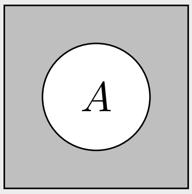

- The *union* of sets $A$ and $B$, written $A \cup B$, is the set of all elements in either $A$ or $B$ or both; *i.e.*,

  $$A \cup B = \{x \in \Omega \mid x \in A \text{ or } x \in B\}$$




- The *complement* of a set $A$, written $\overline{A}$, is the set of all elements in $\Omega$ that are not in $A$; *i.e.*,

  $$\overline{A} = \{x \in \Omega \mid x \notin A\}$$



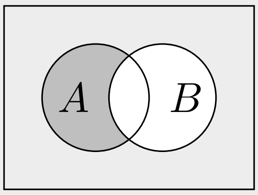





## Set operations

**Definition**

The *set difference* of $A$ and $B$, written $A \setminus B$, is the set of all elements in $A$ but not in $B$; *i.e.*, $A \setminus B = \{x \in \Omega \mid x \in A \;\; \& \;\; x \notin B\}$.




Note that $\overline{A} = \Omega \setminus A$.

**Exercise**

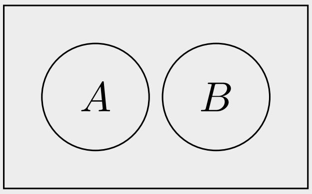
Show that:

- $A \setminus B = A \cap \overline{B}$
- $A \cap \Omega = A$
- $A \cap \varnothing = \varnothing$
- $A(B \cup C) = (AB) \cup (AC)$
- $\overline{A \cup B} = \overline{A} \cap \overline{B}$
- $\overline{A \cap B} = \overline{A} \cup \overline{B}$
- $A \cup \Omega = \Omega$
- $A \cup \varnothing = \varnothing$

# Set defintions

**Definition**

Sets $A$ and $B$ are *disjoint* iff $A \cap B = \varnothing$.




**Exercise**

Show that for sets $A$ and $B$:

- they are disjoint (*i.e.*, $A \cap B = \varnothing$) iff $A \subseteq \overline{B}$
- $A \subseteq B$ iff $A$ and $\overline{B}$ are disjoint
- Show that $A$ and $\overline{A}$ are disjoint
- $A \setminus B$, $AB$, and $B \setminus A$ are all (pair-wise) disjoint





# Cardinality and counting

**Definition (Cardinality)**

The *cardinality* of a set $A$, written $|A|$, is the *magnitude* of $A$; for finite $A$ this is the number of elements of $A$.

Examples: if $A = \{2, 4, 7\}$, then $|A| = 3$. The set of positive square numbers less than $20$ has cardinaity $4$.

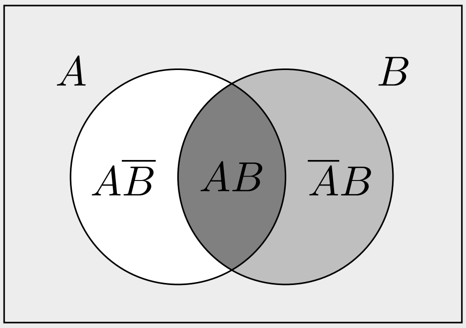
**Definition**

Sets $A$ and $B$ are said the be *equinumerous*, written $A \cong B$, iff there is a one-to-one correspondence from $A$ to $B$. By definition $|A| = |B|$ iff $A \cong B$.

**Exercise**

What is $|\varnothing|$? $|\{n \in \mathbb{N} \mid 0 < n^2 < 2013 \;\;\&\;\; n \text{ is even}\}|$? Show that $\mathbb{N} \cong \{1, 4, 9, 16, \ldots\}$.

# Counting




### Cardinal properties

- If $A$ and $B$ are disjoint, then $|A \cup B| = |A| + |B|$ (Sum rule)
- $|A \times B| = |A| \times |B|$ (Product rule)

### Exercise

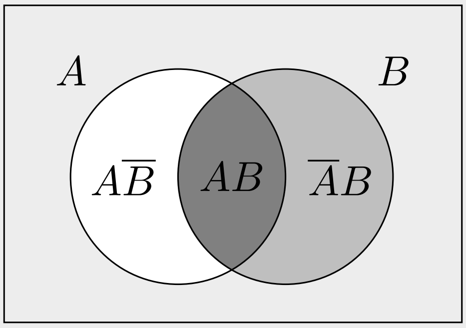
Show that:

- the set of binary sequences of length $n$, $\{0,1\}^n \cong \mathbb{P}(\{1, \ldots, n\})$
- in general, for any $A$ and $B$, $|A \cup B| = |A| + |B| - |A \cap B|$





# Properties




### Exercises

Show that:

- If $B \subseteq A$, then $|A \setminus B| = |A| - |B|$ and $|B| \leqslant |A|$
- Verify that for any set $A$, $|A| \geqslant 0$
- $A \cup B = (A \setminus B) \cup (A \cap B) \cup (B \setminus A) = A\overline{B} \cup AB \cup \overline{A}B$
- What region in the diagram represents $\overline{A} \cap \overline{B}$?
- $A \setminus B$ and $A \cap B$ are disjoint

## Counting example

### Example (Seats)

Seats in a small auditorium are labelled by a row letter $L = \{A, B, C, D, E\}$, and a seat number $N = \{1, \dots, 10\}$.

### Exercise

How many seats are there in total?

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| A | | | | | | | | | | |
| B | | | | | | | | | | |
| C | | | | | | | | | | |
| D | | | | | | | | | | |
| E | | | | | | | | | | |

$$\Omega \cong L \times N$$

$$|\Omega| = |L \times N| = |L| \times |N| = 5 \times 10 = 50.$$





## Probability ontology

### Definition (Conjunction event)

The *conjunction* of the events corresponding to $A, B \subseteq \Omega$ is the event that occurs when both $A$ and $B$ occur simultaneously; *i.e.*, it corresponds to $A \cap B$.

### Definition (Disjunction event)

The *disjunction* of the events corresponding to $A, B \subseteq \Omega$ is the event that occurs when one or both $A$ and $B$ occur; *i.e.*, it corresponds to $A \cup B$.

### Definition (Negation event)

The *negation* of the event corresponding to $A \subseteq \Omega$ is the event that occurs when $A$ and doesn't occur; *i.e.*, it corresponds to $\overline{A}$.

# Probability language

**Definition (Mutually exclusive events)**

The events corresponding to $A, B \subseteq \Omega$ are *mutually exclusive*, or *incompatible*, if $A$ and $B$ are disjoint.

Example: rolling a three an rolling an even number; $T = \{s_3\}$ and $E = \{s_2, s_4, s_6\}$.

**Definition (Impossible and certain events)**

An event is *impossible* if it corresponds to $\varnothing$. An event corresponding to $\Omega$ is *certain*.

**Corollary**

*Two events $A$ and $B$ are incompatible iff their conjunction is impossible.*





# Finite probability

**Definition (Laplace)**

For an experiment whose outcome space $\Omega$ consists of *finitely many equally likely outcomes*, the probability of event $E \subseteq \Omega$ is the ratio of the event's outcomes to the total possible outcomes; *i.e.*:

$$P(E) = \frac{|E|}{|\Omega|}.$$

**Exercise**

What is the probability of 'rolling a six'? Of rolling an even number? A multiple of three? A number greater than two? An even number greater than two? An even number or a number greater than two?

# Finite probability properties

## Exercises

Show that for Laplace's definition of probability:

- $P(\Omega) = 1$
- $P(A) \geqslant 0$, for any event $A \subseteq \Omega$
- if $A$ and $B$ are mutually exclusive events then $P(A \cup B) = P(A) + P(B)$
- $P(\overline{A}) = 1 - P(A)$
- if $A$ and $B$ are not mutually exclusive then $P(A \cup B) = P(A) + P(B) - P(AB)$

If $\Omega$ is not finite, or the possible outcomes are not equally likely, then the probability assignment $P(E) = \frac{|E|}{|\Omega|}$ is no longer meaningful.





# Probability properties

## Exercise

Verify that for any probability function $f$:

- $f(\varnothing) = 0$
- $f(\overline{A}) = 1 - f(A)$
- $f(A \cup B) = f(A) + f(B) - f(AB)$

## Exercise

Verify that for any probability function $f$, if $A \subseteq B$, then $f(A) \leqslant f(B)$.

## Conditional probability

**Example (Dice)**

For a fair die, consider the events:

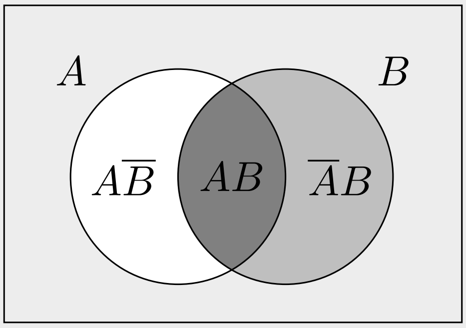
- $E = \{s_2, s_4, s_6\}$—an even number
- $F = \{s_4, s_5, s_6\}$—a number greater than three

- $P(E) = \frac{|E|}{|\Omega|} = \frac{3}{6} = \frac{1}{2}$
- If we learn that a number greater than three was rolled then our new outcome space is $\Omega' = F = \{s_4, s_5, s_6\}$. Moreover, $E' = \{s_4, s_6\}$. So $P_F(E') = \frac{|E'|}{|\Omega'|} = \frac{|\{s_4,s_6\}|}{|\{s_4,s_5,s_6\}|} = \frac{2}{3}$.
  But $E' = E \cap F = EF$ and $\Omega' = \Omega \cap F = F$. Hence $P_F(E') = \frac{P(EF)}{P(F)}$.
- Note here that $P_F(E') > P(E)$





## Conditional probability




**Definition**

The *conditional probability* of event $A$ conditional on $B$ (where $P(B) \neq 0$), written $P(A|B)$, is defined by:

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

In the diagram above, $P(A|B)$ represents the ratio of (the area of) the region $AB$ (the dark region) to that of the whole of $B$.

## Conditional probability

### Exercise

For the dice example earlier:

- What is $P(E|F)$?
- $P(F|E)$?
- Let $G$ be the event: 'The number *three* is rolled'. Describe the event $\overline{G}$ in words
- What are $P(E|\overline{G})$ and $P(\overline{G}|E)$?

### Exercises

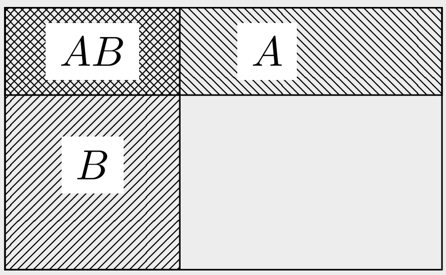
In general, for any events $A, B$:

- Is $P(A|B) = P(B|A)$ necessarily true?
- Is $P(A|B) + P(\overline{A}|B) = 1$?
- Is $P(A|B) + P(A|\overline{B}) = 1$?





## Conditional independence




### Definition

Event $A$ is *(conditionally) independent* of event $B$ if:

$$P(A|B) = P(A).$$

Event $A$ is *(conditionally) dependent* on $B$ if $A$ is not (conditionally) independent of $B$.

For example, if $B$ is a random sample of a population.

## Conditional independence

### Exercises

- If $A$ is conditionally independent of $B$ is $B$ necessarily conditionally independent of $A$?
- Is rolling an even number independent of rolling a number greater than three?
- Prove that if $A$ is conditionally independent of $B$ then $P(AB) = P(A)P(B)$.





## Bayes's rule

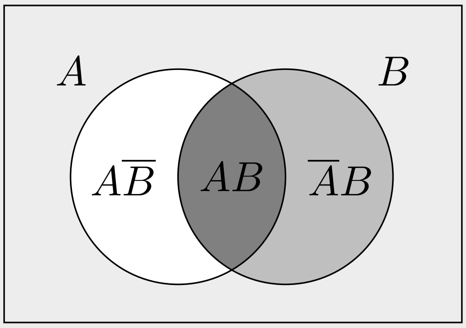
- Rearranging the definition of conditional probability: $P(A \cap B) = P(A|B)P(B)$
- By symmetry $P(A \cap B) = P(B \cap A)$;
  therefore: $P(A|B)P(B) = P(B|A)P(A)$.
  Rearranging this equation gives:

### Theorem (Bayes's Theorem I)

*If $A$ and $B$ are any two events ($P(A) \neq 0$), then:*

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

# Bayes's Venn diagram




But $A = AB \cup A\overline{B}$. So we get the following:

**Theorem (Bayes's Theorem I')**

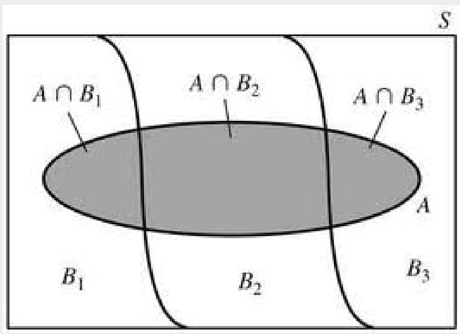
*If $A$ and $B$ are any two events ($P(A) \neq 0$), then:*

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





# Extending Bayes's rule




$$P(B_1|A) = \frac{P(AB_1)}{P(AB_1) + P(AB_2) + P(AB_3)}$$

$$P(B_2|A) = \frac{P(AB_2)}{P(AB_1) + P(AB_2) + P(AB_3)}$$

$$P(B_3|A) = \frac{P(AB_3)}{P(AB_1) + P(AB_2) + P(AB_3)}$$

## Bayes's rule generalised

Events $B_1, \ldots, B_n$ are said to be *universally exhaustive* (of $\Omega$) if $\bigcup_{i=1}^{n} B_i = \Omega$.

### Theorem (Bayes's Theorem II)

*If $B_1, \ldots, B_k, \ldots, B_n$ are mutually exclusive and universally exhaustive events and $A$ is a possible event ($P(A) \neq 0$) then:*

$$P(B_k|A) = \frac{P(AB_k)}{\sum_{i=1}^{n} P(AB_i)}$$

### Theorem (Bayes's Theorem II')

*If $B_1, \ldots, B_k, \ldots, B_n$ are mutually exclusive and universally exhaustive events and $A$ is a possible event ($P(A) \neq 0$) then:*

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





## Bayes's rule example

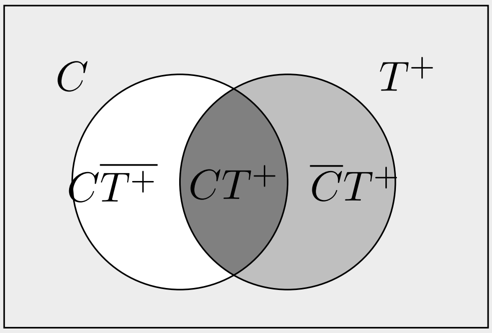
### Example (Medical diagnostics)

In a given population of people, one in every thousand have cancer. A certain pathology test is used to detect the disease. The test is 'good' but not perfect; it returns a positive result for someone with the disease about $98\%$ of the time, and registers a *false positive* (*i.e.*, gives a positive result for a person free of the disease) $5\%$ of the time.

### Exercise

A random person comes in to get tested and the test returns a positive result. What is the probability that the person has cancer?

# Example: solution




Given information:

$$P(C) = \tfrac{1}{1000} \qquad P(\overline{C}) = \tfrac{999}{1000}$$
$$P(T^+|C) = \tfrac{98}{100} \qquad P(\overline{T^+}|C) = \tfrac{2}{100}$$
$$P(T^+|\overline{C}) = \tfrac{5}{100} \qquad P(\overline{T^+}|\overline{C}) = \tfrac{95}{100}$$

Calculate $P(C|T^+)$.

$$\begin{aligned}
P(C|T^+) &= \frac{P(CT^+)}{P(T^+)} = \frac{P(CT^+)}{P(CT^+ \cup \overline{C}T^+)} = \frac{P(CT^+)}{P(CT^+) + P(\overline{C}T^+)} \\
&= \frac{P(T^+|C)P(C)}{P(T^+|C)P(C) + P(T^+|\overline{C})P(\overline{C})} \\
&= \frac{\frac{98}{100} \times \frac{1}{1000}}{\frac{98}{100} \times \frac{1}{1000} + \frac{5}{100} \times \frac{999}{1000}} = \frac{98}{98 + 5 \times 999} \approx \frac{100}{5000} = 0.02
\end{aligned}$$

Patient only has $2\%$ chance of having cancer despite testing positive?!





# Expected value

### Example (Three coins)

Consider the experiment of tossing three coins simultaneously; *i.e.*, $\Omega = \{h,t\} \times \{h,t\} \times \{h,t\} = \{h,t\}^3 = \{hhh, hht, hth, htt, thh, tht, tth, ttt\}$.
Let $N_H(\omega)$ be the number of heads in outcome $\omega$; *e.g.*,

$$N_H(hth) = N_H(hht) = N_H(thh) = 2;$$
$$N_H(ttt) = 0 \qquad N_H(hhh) = 3.$$

In general $N_H(\omega) \in \{0, \ldots, 3\}$. Let $X_k$ be the event of tossing $k$ heads.

$$X_0 = \{ttt\}, \; X_1 = \{htt, tht, tth\},$$
$$X_2 = \{hht, hth, thh\}, \; X_3 = \{hhh\}.$$

Hence:

$$\begin{aligned}
E(h) &= \tfrac{1}{8}(0) + \tfrac{1}{8}(1) + \tfrac{1}{8}(1) + \tfrac{1}{8}(1) + \tfrac{1}{8}(2) + \tfrac{1}{8}(2) + \tfrac{1}{8}(2) + \tfrac{1}{8}(3) \\
&= \tfrac{1}{8}(0) + \tfrac{3}{8}(1) + \tfrac{3}{8}(2) + \tfrac{1}{8}(3) \\
&= P(X_0)(0) + P(X_1)(1) + P(X_2)(2) + P(X_3)(3) \\
&= \tfrac{0+3+6+3}{8} = \tfrac{12}{8} = \tfrac{3}{2}.
\end{aligned}$$

## Expected values

### Definition (Expected value)

The *expected value* of a random variable $X : \Omega \to \mathbb{R}$ with probability distribution $P : \Omega \to \mathbb{R}$ is given by:

$$E(X) = \sum_{\omega \in \Omega} P(\omega) X(\omega)$$

### Definition

The event corresponding to value $x \in \mathbb{R}$, denoted $X_x$, is defined as:

$$X_x = X^{-1}[x] = \{\omega \in \Omega \mid X(\omega) = x\}$$

More generally, for $A \subseteq \mathbb{R}$:

$$X_A = X^{-1}[A] = \{\omega \in \Omega \mid X(\omega) \in A\}$$





## Expected values

### Corollary

*If* $\operatorname{ran} X = \{x_1, \ldots, x_n\}$*, and* $X_{x_i} = X^{-1}[x_i] = \{\omega \in \Omega \mid X(\omega) = x_i\}$*, then events* $X_{x_1}, \ldots, X_{x_n}$ *partition* $\Omega$*. It follows that:*

$$E(X) = \sum_{i=1}^{n} P(X_{x_i}) x_i$$

Often $X$ is referred to as a *random variable*, and $X_{x_i}$ is written $X = x_i$; *i.e.*, $P(X = x_i)$.

### Exercise

For the three-coins example, let $X$ map outcomes to the number of heads. What is $X_2$? $X_{\{2,3\}}$?

# Expected values

For a random variable (real-valued function from $\Omega$ to $\mathbb{R}$) $X$:

- $E(X)$ is also called the limiting (or long run) *average* of $X$
- $E(X)$ may not be any actual value in $\operatorname{ran} X$
- $E(X)$ is a measure of the 'centre', or *centroid*, of the values of the outcomes
- Natural correspondence with the 'centre of gravity/mass' of a distribution of point masses on a line, where $P(X = x_i)$ corresponds to the proportion of the total mass positioned at $x_i$