9. Probability

- Probability: a mathematical model for chance (random) phenomena.
- random process are everywhere:

(1) Weather prediction and climate modelling,
(2) Physics: theory of gases, quantum mechanics
(3) Actuarial science, insurance
(4) Statistics: procedures for analysing data, especially data that has a random character
(5) Genetics and oncology: model for mutations
(6) Transmission of data affected by noise
(7) Games of chance: tossing a coin, rolling a die, card games,

- randomness implies a lack of predictability, we cannot compute the outcome before the event, but we can try to estimate the likelihood of each possible outcome

Definition 9.1. *A* **statistical experiment** *is a process by which we observe something uncertain. An* **outcome** *is a result of a statistical experiment. The set of all outcomes is denoted by* Ω. *An* **event** *is a subset of* Ω. *By* **Events of Interest** *we mean any set* $\mathscr{E}$ *of events, such that*

(E1) $\emptyset \in \mathscr{E}$,
(E2) $\{\omega\} \in \mathscr{E}$, *for all* $\omega \in \Omega$,
(E3) *If* $A \in \mathscr{E}$, *then* $A^c := \Omega \backslash A \in \mathscr{E}$,
(E4) *If* $A_i \in \mathscr{E}$, *for integers* $n \geq 1$, *then* $\bigcup_{i=1}^{\infty} A_i \in \mathscr{E}$

We call the pair $(\Omega, \mathscr{E})$ *the* **sample space** *of the experiment.*

Remark 9.2. *If* Ω *has only finitely many outcomes, then* $\mathscr{E}$ *may contain all subsets of* Ω, *that is,* $\mathscr{E}$. *If* Ω *is infinite, then we have to restrict* $\mathscr{E}$ *to "certain" subsets.*

Example 9.3. (1) *Experiment: "Rolling a die"*
Outcome: "Throwing a six"
Sample space: $\Omega = \{1, 2, 3, 4, 5, 6\}$
Event: $A =$*"Throwing an even number". Then* $A = \{2, 4, 6\}$

(2) *Experiment: "Tossing a coin",*
Sample space: $\Omega = \{heads, tails\}$
Event: $B =$*"Throwing heads". Then* $B = \{heads\}$

(3) *Experiment: "Picking from a full stack of cards"*
Sample space: $\Omega = \{$*2 of spades, 3 of spades*$, \ldots,$ *ace of spades,*
2 of hearts$, \ldots,$ *ace of clubs*$\}$
Event: $C =$*"Picking hearts". Then* $C = \{$*2 of hearts*$, \ldots,$ *ace of hearts*$\}$

(4) *Experiment: "Picking two balls from an urn with red and blue balls"*
Sample space: $\Omega = \{RR, RB, BR, BB\}$
Event: $D =$*"Picking two different colours". Then* $D = \{RB, BR\}$
Event: $E =$*"Picking blue the second time". Then* $D = \{RB, BB\}$
Event: $F = D \cap E = \{RB\}$

(5) *Experiment: "Cycling through 3 traffic lights, which are red or green"*
Sample space: $\Omega = \{RRR, RRG, RGR, GRR, RGG, GRG, GGR, GGG\}$
Event: $G =$*"Stopping at most once. Then* $G = \{RGG, GRG, GGR, GGG\}$

We want to be able to talk about the relative likelihood of "events of interest" in a sample space.

Definition 9.4. *Let* $(\Omega, \mathscr{E})$ *be the sample space of an experiment. A* **probability measure** *is a function* $P : \mathscr{E} \to \mathbb{R}$*, such that the following* **axioms of probability** *hold:*

(AP1) $P(\Omega) = 1$*;*
(AP2) *If* $A \in \mathscr{E}$ *then* $P(A) \geq 0$*;*
(AP3) *If* A *and* B *are disjoint events in* $\mathscr{E}$*, then* $P(A \cup B) = P(A) + P(B)$.

For any outcome $\omega \in \Omega$*, we set* $P(\omega) := P(\{\omega\})$*. We call the triplet* $(\Omega, \mathscr{E}, P)$ *a* **probability space***.*

Lemma 9.5. *Let* $(\Omega, \mathscr{E}, P)$ *be a probability space of an experiment and let* $A, B \in \mathscr{E}$ *be events. Then*

(1) $P(A^c) = 1 - P(A)$
(2) $P(\emptyset) = 0$
(3) *If* $A \subseteq B$ *then* $P(A) \leq P(B)$.
(4) $P(A) \in [0, 1]$
(5) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
(6) *If* $A = \{\omega_1, \ldots, \omega_n\}$*, then* $P(A) = P(\omega_1) + \cdots + P(\omega_n)$

Proof. (1) We have that A and A^c are disjoint and $\Omega = A \cup A^c$. Hence by (PA1) and (PA3) we get

$$1 = P(\Omega) = P(A \cup A^c) = P(A) + P(A^c).$$

(2) As $\emptyset = \Omega^c$, it follows from (1) and (PA1) that $P(\emptyset) = P(\Omega^c) = 1 - P(\Omega) = 0$.

(3) Note that $B = A \cup (B \backslash A)$, where A and $B \backslash A$ are disjoint. Also $P(B \backslash A) \geq 0$, by (PA2). Hence by (PA3) we have

$$P(A) \leq P(A) + P(B \backslash A) = P(A \cup (B \backslash A)) = P(B).$$

(4) We have $P(A) \geq 0$, by (PA2). As $A \subseteq \Omega$, it follows from (3) and (PA1) that $P(A) \leq P(\Omega) = 1$.

(5) Note that $A \cup B = A \cup (B \backslash A)$, where A and $B \backslash A$ are disjoint. Hence, by (PA3),

$$P(A \cup B) = P(A) + P(B \backslash A).$$

Next observe that $B = (A \cap B) \cup (B \backslash A)$, where $(A \cap B)$ and $B \backslash A$ are disjoint. Hence, by (PA3),

$$P(B) = P(A \cap B) + P(B \backslash A).$$

Putting both equations together we get (5).

(6) The sets $\{\omega_1\}, \ldots, \{\omega_n\}$ are pairwise disjoint and thus a repeated application of (PA3) gives the result. $\square$

Lemma 9.6. *Let $\Omega = \{\omega_1, \ldots, \omega_n\}$. Furthermore assume that $P : \Omega \to [0, 1]$, such that*

$$\sum_{i=1}^{n} P(\omega_i) = 1.$$

Then P becomes a probability measure on the sample space $(\Omega, \mathcal{P}(\Omega))$ by setting

$$P(A) = \sum_{\omega \in A} P(\omega)$$

Example 9.7. *(1) A coin is tossed. Hence we have the sample space $(\Omega, \mathcal{P}(\Omega))$, where $\Omega = \{heads, tails\}$. Now $P(heads) = P(tails) = \frac{1}{2}$, gives rise to a probability space $(\Omega, \mathcal{P}(\Omega), P)$. In this case we call the coin* **fair**. *Likewise $Q(heads) = 0.1$ and $Q(tails) = 0.9$ gives rise to a probability space $(\Omega, \mathcal{P}(\Omega), Q)$.*

(2) In a presidential race there are four candidates $C1$, $C2$, $C3$ and $C4$ and polling suggests that their respective chance of winning is 25%, 15%, 30% and 30%. Consider the probability space $(\Omega, \mathcal{P}(\Omega), P)$, where $\Omega = \{C1, C2, C3, C4\}$ and $P(C1) = 0.25$, $P(C2) = 0.15$ and $P(C3) = P(C4) = 0.3$. If A is the event of $C1$ or $C3$ winning, then $P(A) = P(C1) + P(C3) = 0.25 + 0.3 = 0.55$. Note

that $A^c = \{C2, C4\}$ *and so* $P(A^c) = 1 - P(A) = 0.45 = P(C2) + P(C4)$. *If* B *is the event of* $C3$ *or* $C4$ *winning, then* $P(B) = P(C3) + P(C4) = 0.3 + 0.3 = 0.6$. *Note that* $A \cap B = \{C3\}$ *and so* $P(A \cap B) = P(C3) = 0.3$. *Finally,* $A \cup B = \{C1, C3, C4\}$ *and* $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.55 + 0.6 - 0.3 = 0.85 = P(C1) + P(C3) + P(C4)$.

Corollary 9.8. *Let* $(\Omega, \mathcal{P}(\Omega), P)$ *be a finite probability space, where each outcome occurs with the same probability. Then for every event* A

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{number of ways } A \text{ can occur}}{\text{total number of outcomes}}$$

Example 9.9. *(1) Toss a fair coin, with outcomes "H" and "T", twice. Then* $\Omega = \{HH, HT, TH, TT\}$. *Note that* $|\Omega| = 4$ *and so each outcome has probability* 0.25. *Next let* A *be the event of heads on first toss and* B *the event of heads on first or second toss. Then* $A = \{HH, HT\}$, $B = \{HT, TH, HH\}$. *Now*

$$P(A) = \frac{|A|}{|\Omega|} = \frac{2}{4} = 0.5 \quad \text{and} \quad P(B) = \frac{|B|}{|\Omega|} = \frac{3}{4} = 0.75.$$

(2) Two fair dice are rolled in succession. Let A *denote the event of rolling an 8 in total. What is* $P(A)$*? Observe that* $\Omega = \{(i, j) \mid 1 \le i, j \le 6\}$, *where each outcome* (i, j) *is equally likely, that is,* $P(i, j) = \frac{1}{|\Omega|} = \frac{1}{36}$. *Then*

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

and so $P(A) = \frac{|A|}{|\Omega|} = \frac{5}{36}$.