8 Independent Events

Definition We say that the events A and B are *independent* if

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

You may assume that two events are independent in the following situations:

- (i) they are clearly physically unrelated (eg depend on different coin tosses),
- (ii) you calculate their probabilities and find that $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$
- (iii) the question tells you that the events are independent!

Remark (Warning)

(a) Independence is a weaker property than being physically unrelated. For example we saw in lectures that if a fair die is rolled twice then the event "first roll is a 6" and the event "both rolls produce the same number" are independent even though they are physically related.

(b) Saying two events are independent is completely different to saying they are disjoint: two disjoint events are *never* independent (unless one has probability 0). Why not?.

If we have more than two events things get more complicated. We saw an example where any two of A, B, C were independent but

$$\mathbb{P}(A \cap B \cap C) \neq \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C).$$

Definition We say that the events A_1, A_2, \ldots, A_n are mutually independent if for all $1 \le i_1 < i_2 < \cdots < i_t \le n$ we have

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_t}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2}) \dots \mathbb{P}(A_{i_t}).$$

That is to say for every subset I of the events, the probability that all events in I occur is the product of the probabilities of the individual events in I. If you find this confusing then think what it says in the case n = 3 first.

Example A coin which has probability p of showing heads is tossed n times in succession. What is the probability that heads comes up exactly r times?

We may assume that the results of the tosses are mutually independent (they are physically unrelated things). So to find the probability of a particular sequence of r heads and n - r tails we just multiply together the appropriate probability for each toss. That is

$$\mathbb{P}(\underbrace{h\dots h}_{r \text{ heads}} \underbrace{t\dots t}_{n-r \text{ tails}}) = p^r (1-p)^{n-r}.$$

Any other sequence of r heads and n-r tails will have the same probability $p^r(1-p)^{n-r}$. There are $\binom{n}{r}$ such sequences and so

$$\mathbb{P}(r \text{ heads and } n-r \text{ tails in any order}) = \binom{n}{r} p^r (1-p)^{n-r}$$

We will see this example again when we discuss the binomial distribution.

We saw several examples of calculating probabilities of events like this i.e. events which are intersections and unions of several independent events. There are more examples on the exercise sheets.

9 Conditional Probability

Suppose A and B are events in a sample space S. If we are told that B has occurred what effect does this have on the probability that A has occurred?

Definition If A and B are events and $\mathbb{P}(B) \neq 0$ then the *conditional* probability of A given B is

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

This is usually denoted by $\mathbb{P}(A|B)$.

Note that the definition does not assume that A happens after B. One way of thinking of this is to imagine the experiment is performed secretly and the fact that B occurred is revealed to you (without the full outcome being revealed). The conditional probability of A given B is the new probability of A in these circumstances.

This is an important definition but it is often the source of great confusion; make sure you understand it and can use it. Also, don't confuse $\mathbb{P}(A|B)$ with $\mathbb{P}(A \setminus B)$.

One use of conditional probability is to calculate the probability of the intersection of several events. We saw in the last section that if the events are mutually independent then the probability of their intersection is just the product of their probabilities. If the events are not mutually independent then we need to use conditional probability. We will need to consider conditional probabilities like $\mathbb{P}(A_3|A_1 \cap A_2)$ i.e. the conditional probability of A_1 given that both A_2 and A_3 occur.

Theorem 9.1. Let A_1, A_2, \ldots, A_n be events in a sample space S. Suppose that $\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_{n-1}) > 0$. Then

$$\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_n) = \\\mathbb{P}(A_1) \times \mathbb{P}(A_2 | A_1) \times \mathbb{P}(A_3 | A_1 \cap A_2) \times \ldots \times \mathbb{P}(A_n | A_1 \cap A_2 \cap \ldots \cap A_{n-1}).$$

We use induction on n to prove this. The idea is that we first prove that the statement is true for n = 2 (the base case of the induction). We then assume that $n \ge 3$ and that the statement of the theorem is true whenever we have less than n events (the induction hypothesis). We then use this assumption to prove that the statement of the theorem is true when we have n events (the inductive step). This is an important proof technique. You will meet it again in many other modules.

Proof of Theorem 9.1

Base case. Suppose n = 2. Using the definition of conditional probability we have

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

since $A_2 \cap A_1 = A_1 \cap A_2$. Thus the statement of the theorem is true when n = 2.

Induction Hypothesis Suppose that $n \geq 3$ and that the statement of the theorem is true whenever we have less than n events. In particular the statement of the theorem is true when we have n-1 events. Since $\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_{n-2}) \geq \mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_{n-1}) > 0$ we may deduce that

$$\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_{n-1}) = \mathbb{P}(A_1) \times \mathbb{P}(A_2 | A_1) \times \mathbb{P}(A_3 | A_1 \cap A_2) \times \ldots \times \mathbb{P}(A_{n-1} | A_1 \cap A_2 \cap \ldots \cap A_{n-2}).$$

Inductive Step Let $B = A_1 \cap A_2 \cap \cdots \cap A_{n-1}$. Then

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

$$= \mathbb{P}(B \cap A_n)$$

$$= \mathbb{P}(B) \times \mathbb{P}(A_n | B)$$

$$= \mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \times \mathbb{P}(A_n | A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

$$= \mathbb{P}(A_1) \times \mathbb{P}(A_2 | A_1) \times \mathbb{P}(A_3 | A_1 \cap A_2) \times \cdots \times \mathbb{P}(A_n | A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

where the second equality uses the base case to evaluate $\mathbb{P}(B \cap A_n)$, the third equality is a direct substitution for B, and the fourth equality uses the induction hypothesis to evaluate $\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_{n-1})$. Thus the truth of of the theorem for less than n events implies the theorem is also true when we have n events.

We may now deduce that the theorem is true for all values of n.