Chapter 8

Probability and Statistics

8.1 Basic Probability

For an event E, the probability of the E occurring, denoted $\mathrm{P}(E)$, is a number such that

$$0 \leq \mathrm{P}(E) \leq 1.$$

where

$$\begin{aligned} \mathrm{P}(E) &= 0 \implies E \text{ is impossible,} \\ \mathrm{P}(E) &= 1 \implies E \text{ is certain.} \end{aligned}$$

Example 8.1 (Rolling a die)**.** The set of all possible outcomes is the *sample space*, denoted S, i.e.

$$S = \{1, 2, 3, 4, 5, 6\}$$

Let A be the event of getting an even number in one roll. Then we have

$$A = \{2, 4, 6\}$$

and therefore

$$\mathrm{P}(A) = \frac{3}{6} = \frac{1}{2}.$$

Example 8.2. We randomly select 2 lightbulbs from a set of 5 bulbs (numbered 1 to 5). The sample space consists of 10 possible outcomes:

$$\begin{aligned} S = \; &\{\{1,2\}, \{1,3\}, \{1,4\}, \{1,5\}, \{2,3\}, \\ &\{2,4\}, \{2,5\}, \{3,4\}, \{3,5\}, \{4,5\}\}. \end{aligned}$$

Note that $|S| = 10$ is the number of elements in S, also known as the *cardinality* of the set S. We may be interested in the following events:

A: No faulty bulbs

B: One faulty bulb

C: Two faulty bulbs

Now assume that bulbs 1, 2 and 3 are all faulty. We see that event A occurs only if we draw bulbs 4 and 5 (i.e. outcome $\{4,5\}$).

$$\therefore \quad \mathrm{P}(A) = \frac{1}{10}.$$

Event B occurs if we draw $\{1,4\},\{1,5\},\{2,4\},\{2,5\},\{3,4\}$ or $\{3,5\}$. Hence

$$\mathrm{P}(B) = \frac{6}{10}.$$

Meanwhile, Event C occurs if we draw $\{1,2\},\{1,3\},\{2,3\}$, and therefore

$$\mathrm{P}(C) = \frac{3}{10}.$$

Definition 8.1. *The set of all elements (outcomes) not in E in the sample space S is called the* compliment *of E, usually denoted E^c or $\bar{E}$.*

Example 8.3. E = randomly rolled die gives an even number, i.e.

$$E = \{2,4,6\}$$

then E^c = randomly rolled die gives an odd number, i.e.

$$E^c = \{1,3,5\}$$

Let A and B be two events in an experiment.

Definition 8.2. *The event consisting of all the elements of the sample space that belong to either A or B is called* the union *of A and B and is denoted as $A \cup B$.*

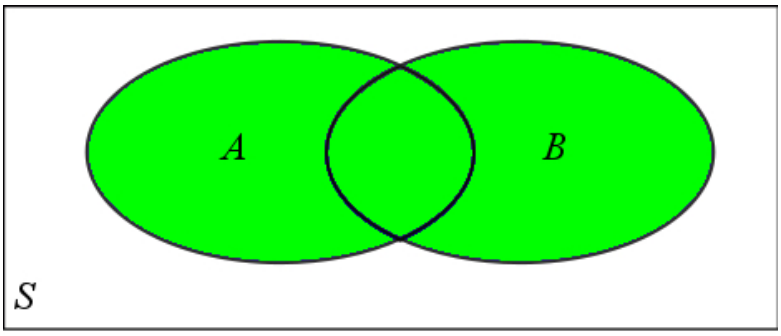

Figure 8.1: A Venn diagram. The union $A \cap B$ is shaded in green.

Definition 8.3. *The event consisting of all the elements of the sample space that belong to both A and B is called* the intersection *of A and B and is denoted as $A \cap B$.*

Example 8.4. Suppose that we are rolling a die, then consider the following events:

A: The die gives a number not smaller than 4.

B: The die gives a number that is a multiple of 3

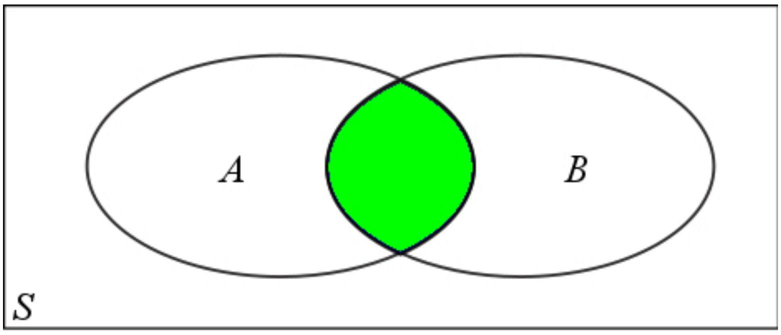

Figure 8.2: A Venn diagram. The intersection $A \cup B$ is shaded in green.

$$A = \{4, 5, 6\}, \quad B = \{3, 6\},$$

then

$$A \cup B = \{3, 4, 5, 6\}, \quad A \cap B = \{6\}.$$

Definition 8.4. *Events A and B are said to be* mutually exclusive *events if they have no element in common, i.e. if*

$$A \cap B = \{\} = \emptyset,$$

where the symbol $\emptyset$ denotes the empty set*. It has no elements, so the cardinality of the empty set is zero.*

The Axioms of Probability

1. For any event E in a sample space S,

$$0 \leq \mathrm{P}(E) \leq 1.$$

2. For the entire sample space S, we have $\mathrm{P}(S) = 1$.

3. If A and B are mutually exclusive events, then

$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B).$$

Fact: If A and B are any events, then

$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B).$$

Fact:

$$\mathrm{P}(E) = 1 - \mathrm{P}(E^c).$$

i.e. the probability of E occurring is 1 $-$ (the probability of E not occurring).

Example 8.5 (Rolling a die again!)**.** The event space is

$$S = \{1, 2, 3, 4, 5, 6\}$$

with $\mathrm{P}(1) = 1/6$, $\mathrm{P}(2) = 1/6$, etc.

A: The event that an even number is given.

$$P(A) = P(2) + P(4) + P(6) = \frac{1}{2}.$$

B: The event that a number greater than 4 turns up.

$$P(B) = P(5) + P(6) = \frac{1}{3}.$$

Example 8.6. Five coins are tossed simultaneously. What is the probability of obtaining at least one head?

Note: There are in total $2^5 = 32$ possible outcomes, only one of which has no heads. Therefore

$$\begin{aligned} P(\text{At Least One Head}) &= 1 - P(\text{No Heads}) \\ &= 1 - \frac{1}{32} = \frac{31}{32}. \end{aligned}$$

Example 8.7. The probability that a person watches TV is $P(T) = 0.6$; the probability that the same person listens to the radio $P(R) = 0.3$. The probability that they do both is 0.15. What is the probability that they do neither?

Using the addition law,

$$\begin{aligned} P(T \cup R) &= P(T) + P(R) - P(T \cap R) \\ &= 0.6 + 0.3 - 0.15 = 0.75. \end{aligned}$$

$$\therefore \quad P(\text{They do neither}) = 1 - P(T \cup R) = 0.25.$$

Conditional probability

Often it is required to find the probability of an event B given that an event A has already occurred. This is known as the *conditional probability* of B given A, and is denoted $P(B|A)$.

The intuition behind this is that A gives a "reduced sample space", and therefore

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

Example 8.8 (Conditional Probability)**.** The probability $P(A)$ that it rains in Manchester on July 15^{th} is 0.6, while the probability $P(A \cap B)$ that it rains there on both the 15^{th} and 16^{th} is 0.35. Given that it rains on the 15^{th}, what is the probability that it rains on the next day?

Note: B is the event that it rains in Manchester on July 16^{th}. We need to find $P(B|A)$, and using the formula for conditional probability :

$$P(B|A) = \frac{P\, A \cap B}{P(A)} = \frac{0.35}{0.6} = \frac{7}{12} = 0.583. \quad (3\text{ d.p})$$

Example 8.9. A fridge contains 10 cans of lager, three of which are "4X" (to be avoided). Robb selects 2 cans at random. Find the probability that none of the selected cans are "4X".

$$\text{Let} \quad A = \text{First can selected is not 4X,}$$
$$B = \text{Second can selected is not 4X.}$$

We will look at two different cases...

1 The case with replacement, i.e. Robb puts the first can back in the fridge before choosing the second. Then
$$\mathrm{P}(A) = \mathrm{P}(B) = \frac{7}{10},$$
and
$$\mathrm{P}(A \cap B) = \frac{7}{10} \times \frac{7}{10} = 0.49.$$

2 Sampling without replacement, i.e. the first can is NOT put back in the fridge. Then...
$$\mathrm{P}(A) = \frac{7}{10}, \quad \text{and} \quad \mathrm{P}(B|A) = \frac{7 \times 1}{10 \times 1} = \frac{6}{9} = \frac{2}{3}.$$
$$\therefore \quad \mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B|A) = \frac{7}{10} \times \frac{2}{3} = \frac{14}{30} \approx 0.47.$$

8.2 Random Variables

Sometimes engineers must work with a variable X whose (real) value is subject to variations due to chance (randomness). We call X a *random variable.*

So X can take on a set of possible different values, each with a corresponding probability. We can say that for each possible value a, for

$$X = a \quad \text{the probability of this value is} \quad \mathrm{P}(X = a).$$

We can then say that the probability that X assumes any value within the range:

1. $b < X < c$ is $\mathrm{P}(b < X < c)$
2. $X \leq c$ is $\mathrm{P}(X \leq c)$
3. $X > c$ is $\mathrm{P}(X > c)$.

Actually,
$$\mathrm{P}(X \leq c) + \mathrm{P}(X > c) = \mathrm{P}(\text{All possible values of } X) = 1,$$
or equivalently,
$$\mathrm{P}(X > c) = 1 - \mathrm{P}(X \leq c).$$

Example 8.10. Let

$$X = \text{Score obtained when I roll a fair die..}$$

Then...

$$\mathrm{P}(X = 1) = \frac{1}{6}, \qquad \mathrm{P}(1 \leq X \leq 2) = \frac{1}{3}$$
$$\mathrm{P}(1 < X < 2) = 0, \qquad \mathrm{P}(X < 0.5) = 0.$$

In this example, our random variable is *discrete.* Random variables can also be continuous, but we will only discuss discrete ones in this course.

Let $x_1, x_2, \ldots$ be the possible values of X, each with probabilities $\mathrm{P}_1, \mathrm{P}_2, \ldots$

Then we can consider a *probability distribution function* (p.d.f) for $f(x)$.

Note that the condition $\sum_j f(x_j) = \sum_j \mathrm{P}_j = 1$ is necessary.

Example 8.11 (Rolling one die)**.** By sketching the p.d.f, we can visualise the distribution of the random variable X...

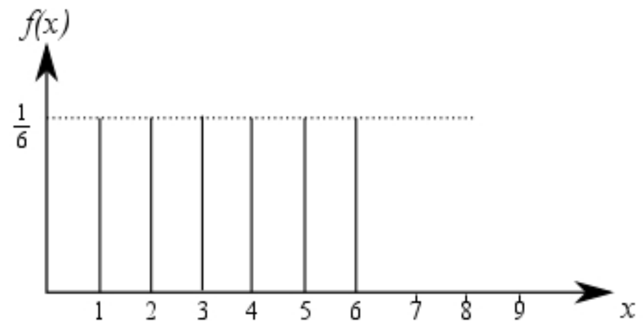

Figure 8.3: The p.d.f. for rolling one die. Observe that the probabilities for the scores 1 to 6 are all the same (and add up to one). Moreover, the p.d.f. shows that there is no chance of scoring 7, 8, 9, ...

This particular example is a uniformly distributed random variable.

Example 8.12 (Rolling two dice)**.** There are 36 possible outcomes, all with a probability of $\frac{1}{36}$. Let's define the random variable X as:

$$X = \text{Sum of the numbers obtained by rolling two dice.}$$

x	2	3	4	5	6	7	8	9	10	11	12
$f(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Example 8.13. Suppose $X = \{0, 1, 2, 3\}$, and the following two distributions are:

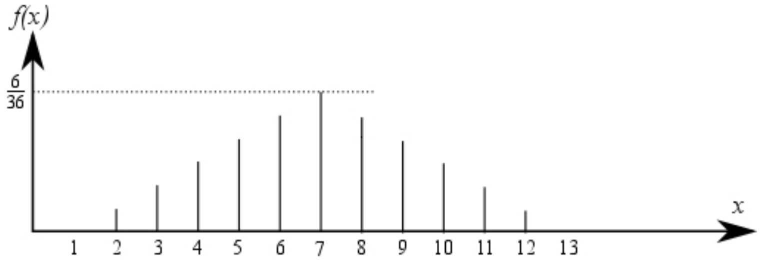

Figure 8.4: The p.d.f. for rolling two dice. Unsurprisingly, there is zero chance of gaining a sum of thirteen!

i $f(x) = \frac{1}{8}(1+x)$

ii $f(x) = \frac{1}{10}(1+x)$.

Only one of these is a valid p.d.f. Which one, and why?

Answer: (ii) is valid, but (i) is not.
Why: Need $\sum_j f(x_j) = 1$. Only (ii) satisfies it.

Definition The *mean, expectation or expected value* μ of a discrete p.d.f:

$$[(\mathrm{E}(X) =] \quad \mu = \sum_j x_j f(x_j) = x_1 f(x_1) + x_2 f(x_2) + \cdots .$$

Example 8.14 (Expected value for rolling a fair die)**.** Recall that

$$f(x_j) = \frac{1}{6} \quad \text{when} \quad j = 1, 2, \ldots, 6$$

$$\Rightarrow \quad \mu = 1 \times \frac{1}{6} + 2 \times \frac{2}{6} + \ldots + 6 \times \frac{6}{6} = 3.5.$$

Granted, we can't gain a score of 3.5 if we roll the die only once. But that is not what μ means. Actually, μ represents the average "score" you would get if you rolled the die many times.

Example 8.15. A stranger shows you a game where you draw a ball out of a bag. There are 6 white balls and 4 blue balls in the bag.

- If the ball is white, you win 40p.
- If the ball is blue, you lose 80p.

Afterwards, the ball is replaced. What are your expected winnings? And is it worth playing that game?

Let X = winnings obtained after drawing the ball out.

$$\text{When } X = 40 \quad (x_1) \quad \mathrm{P}(x_1) = \frac{6}{10},$$
$$X = -80 \quad (x_2) \quad \mathrm{P}(x_2) = \frac{4}{10}.$$

Therefore for the expected value

$$\Rightarrow \qquad \mu = x_1\,\mathrm{P}(x_1) + x_2\,\mathrm{P}(x_2) = 40 \times \frac{6}{10} + (-80) \times \frac{4}{10} = -8\text{p}.$$

$\therefore$ After playing n games you can expect to lose $8n$ pence!

Better off to NOT play this game.

Definition The *variance* of a distribution, denoted σ^2 (or $\mathrm{Var}(X)$) is defined by

$$\begin{aligned}\sigma^2 = \mathrm{Var}(X) &= \sum_j (x_j - \mu)^2 f(x_j)\\ &= (x_1 - \mu)^2 f(x_1) + (x_2 - \mu)^2 f(x_2) + \cdots .\end{aligned}$$

Shortcut: $\sigma^2 = \mathrm{E}(X^2) - \mu^2$, where $\mathrm{E}(X^2)$ is the mean for X^2.

$$\left[\mathrm{E}(X^2) = \sum_j f(x_j) x_j^2.\right]$$

Can interpret σ^2 as a measure of the spread of the data. Specifically, it is the expected square deviation of X from the mean μ.

Example 8.16 (Coin toss)**.** Let 1 and 0 denote heads and tails respectively. It is easy to show that

$$\mu = 0 \times \frac{1}{2} + 1 \times \frac{1}{2} = \frac{1}{2},$$

but what is the variance?
Take the shortcut...

$$\sigma^2 = \left(0^2 \times \frac{1}{2} + 1^2 \times \frac{1}{2}\right) - \left(\frac{1}{2}\right)^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}.$$

8.3 The Binomial Distribution

Start by conducting an experiment (trial) with only two outcomes. They can be labelled "success" or "failure", and their repective probabilities are p and $q = 1 - p$.

E.g. Scoring a 6 from a die roll: $p = \frac{1}{6}$, $q = \frac{5}{6}$.

Then if the trial is repeated a fixed number of times (n), define a new discrete random variable:

$$X = \text{Number of successes in } n \text{ trials.}$$

We assume four conditions:

1. The trial must only have <u>two</u> outcomes

2. Fixed number of trials
3. The probability of success must be the same for all trials
4. The trials are independent.

Example 8.17. Find the probability of 0,1,2,3,4 successes in an experiment consisting of up to 4 repeated trials with probability of success p ($\therefore\ q = 1 - p$).

Number of Trials Number of Successes	1	2	3	4
0	q	q^2	q^3	q^4
1	p	$2pq$	$3pq^2$	$4pq^3$
2	0	p^2	$3p^2q$	$6p^2q^2$
3	0	0	p^3	$4p^3q$
4	0	0	0	p^4

Generally, we can consider the p.d.f. $f(x) = \mathrm{P}(X = x)$. Then the probability of x successes in n trials is

$$\mathrm{P}(X = x) = \binom{n}{x} p^x q^{n-x},$$

where $\binom{n}{x}$ is the binomial coefficient, and the p.d.f. corresponds to the *Binomial Distribution.*

Recall that

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

These binomial coefficients represent the number of ways of choosing x objects from a set of n objects.

Example 8.18. We roll a die 56 times. What is the probability of getting at least three sixes?

Define a random variable X as

$$X = \text{Number of sixes thrown in 56 trials.}$$

Then we say that

$$X \sim \mathrm{B}\left(n = 56, p = \frac{1}{6}\right)$$

Then we want

$$\mathrm{P}(\geq 3) = 1 - \mathrm{P}(X = 0, 1 \,\text{or}\, 2)$$
$$1 - \left[\left(\frac{5}{6}\right)^{56} + \binom{56}{1}\left(\frac{5}{6}\right)^{55}\left(\frac{1}{6}\right) + \binom{56}{2}\left(\frac{5}{6}\right)^{54}\left(\frac{1}{6}\right)^{2}\right]$$

Note: It is perfectly fine to leave your answer in this form!

Example 8.19. A factory produces plenty of board pens. However, 10% of the pens are defective. If I open a random box containing twenty board pens, what is the probability that:

i Exactly 3 pens are defective?

ii More than 3 pens are defective?

(Answer to 3 decimal places)

First, if X = number of faulty pens in a box of 20,

$$X \sim \mathrm{B}(20, 0.1)$$

i We want
$$\mathrm{P}(X = 3) = \binom{20}{3}(0.1)^3(0.9)^{17} \approx 0.190.$$

ii This is $\mathrm{P}(X \geq 3)$, i.e.
$$\begin{aligned}\mathrm{P}(X \geq 3) &= 1 - \mathrm{P}(X \leq 2) \\ &= 1 - \left[0.9^{20} + \binom{20}{1}(0.1)(0.9)^{19} + \binom{20}{2}(0.1)^2(0.9)^{18}\right] \\ &\approx 0.323.\end{aligned}$$

Mean and variance of $\mathrm{B}(n, p)$

Since
$$f(x) = \binom{n}{x} p^x q^{1-x},$$
it turns out that
$$\begin{aligned}\text{Mean:} \quad & \mu = \sum_{x=0}^{n} x f(x) = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} x = np \\ \text{Variance:} \quad & \sigma^2 = npq = np(1-p).\end{aligned}$$

So for the board pen example, $\mu = 2$, $\sigma^2 = 1.8$.

8.4 The Poisson Distribution

Consider the following scenarios:

i Number of phone calls arriving at a call centre per hour.

ii Number of cars crossing a bridge per hour.

iii Number of faults in a length of cable.

These problems require a distribution that involves an average rate μ. Actually, there is one - it is the *Poisson distribution*, and its p.d.f. is:

$$\mathrm{P}(X = x) = \frac{e^{-\mu}\mu^x}{x!},$$

where $X = 0, 1, 2, \ldots,$ to ∞.

Example 8.20. On average, 240 cars per hour pass a check point, and a queue forms if more than three cars pass through in a given minute.

What is the probability of a queue forming in a randomly selected minute?

$$\text{Average number of cars per minute} = \frac{240}{60} = 4 = \mu.$$

Let

$$X = \text{Number of cars passing at a randomly selected minute.}$$

Then $X \sim \mathrm{Po}(4)$, and we require

$$\begin{aligned}\mathrm{P}(X \geq 3) = 1 - \mathrm{P}(0 \leq X \leq 3)&\\ 1 - [\mathrm{P}(X = 0) + \mathrm{P}(X = 1) + \mathrm{P}(X = 2) + \mathrm{P}(X = 3)]&\\ = 1 - 0.4331 = 0.5669.&\end{aligned}$$

One important use of the Poisson distribution is to APPROXIMATE the Binomial distribution, because Poisson is easier to compute.

Recall that for binomial,

$$f(x) = \binom{n}{x} p^x q^{n-x}.$$

Then if you let $p \longrightarrow 0$ and $n \longrightarrow \infty$ with $\mu = np$ fixed and finite,

$$f(x) \longrightarrow \mathrm{Po}(\mu).$$

Moreover, the Poisson distribution has mean μ and variance μ.

Example 8.21. A factory produces screws. The probability that a randomly selected screw is defective is given by $p = 0.01$.

In a random sample of 100 screws, what is the probability that there will be more than two defective screws?

$$\begin{aligned}\text{Let} \quad & A = \text{More than two defective screws}\\ \Rightarrow \quad & A^C = \text{At most 1 defective.}\end{aligned}$$

$$\begin{aligned}\mathrm{P}(A^C) = &\binom{100}{0}(0.01)^0(0.99)^{100} + \binom{100}{1}(0.01)^1(0.99)^{99}\\ &+ \binom{100}{2}(0.01)^2(0.99)^{98}.\end{aligned}$$

After spending ages on your calculator, you finally get

$$\Rightarrow \quad \mathrm{P}(A) = 1 - \mathrm{P}(A^C) \approx 0.0794. \quad (3\,\text{s.f.})$$

Alternative: Poisson approximation. As n is large and p small, we have

$$\mu = np = 1, \qquad \therefore \text{ 1 out of 100 defective on average.}$$

$$\Rightarrow \quad \mathrm{P}(A^C) \approx e^{-1}\left(\frac{1^0}{0!} + \frac{1^1}{1!} + \frac{1^2}{2!}\right) = \times\frac{5}{2}e^{-1} \approx 0.9197,$$

and

$$\mathrm{P}(A) = 1 - \mathrm{P}(A^C) \approx 0.0803. \qquad \text{Close to the binomial result!}$$