2 The Poisson distribution

This chapter will show you how to
- identify real-life situations which can be modelled by the Poisson distribution
- use probability tables for the Poisson distribution where available
- calculate probabilities for a general Poisson distribution
- work with the mean and variance of the Poisson distribution
- approximate a binomial distribution by a Poisson distribution.

Before you start

You should know how to:

1 Use your calculator to work out values of exponential functions.

2 Evaluate expressions involving the exponential function.

Check in:

1 Find the value of
  a e^{-3}
  b $e^{-2.1}$

2 Find the value of
$$p = \frac{e^{-3} \times 3^5}{5!}$$

S2

2.1 Introducing the Poisson distribution

Think about the following random variables:

- the number of dandelions in a square metre of open ground
- the number of errors in a page of a typed manuscript
- the number of cars passing under a bridge on a motorway in a minute (when there is no traffic interference on the motorway)
- the number of telephone calls received by a company switchboard in half an hour.

Do these random variables have any features in common?

The behaviour of these random variables follows the **Poisson distribution**.

Formally, the conditions for a Poisson distribution are that:

1 events occur at random
2 events occur independently of one another
3 the average rate of occurrences remains constant
4 there is zero probability of simultaneous occurrences.

The Poisson distribution is defined as

$$P(X = r) = \frac{e^{-\lambda}\lambda^r}{r!} \quad \text{for } r = 0, 1, 2, 3, \ldots$$

There is a **family** of Poisson distributions with only one parameter, λ, which is the mean number of occurrences in the time period (or length, or area) being considered.

You can write the Poisson distribution as $X \sim \text{Po}(\lambda)$.

EXAMPLE 1

If $X \sim \text{Po}(3)$, find $P(X = 2)$.

$$P(X = 2) = \frac{e^{-3}3^2}{2!}$$
$$= 0.224 \text{ (3 s.f.)}$$

EXAMPLE 2

The number of cars passing a point on a road during a 5-minute period may be modelled by the Poisson distribution with parameter 4.
Find the probability that in a 5-minute period

a 2 cars go past **b** fewer than 3 cars go past.

$X \sim \text{Po}(4)$

a $P(X = 2) = \dfrac{e^{-4}4^2}{2!} = 0.146\,525\ldots = 0.147$ (3 s.f.)

b $P(X = 0) = \dfrac{e^{-4}4^0}{0!} = 0.01\,831\ldots = 0.0183$ (3 s.f.)

$P(X = 1) = \dfrac{e^{-4}4^1}{1!} = 0.07\,326\ldots = 0.0733$ (3 s.f.)

$P(X \leqslant 2) = 0.01\,831\ldots + 0.07\,326\ldots + 0.146\,525\ldots$
$= 0.238$ (3 s.f.)

The formulae and tables booklet has the probabilities for some values of λ, but not all, so you need to be able to use the formula as well. Practise using it to improve your confidence and accuracy.

S2

Exercise 2.1

1 If $X \sim \text{Po}(2)$ find **a** $P(X = 1)$ **b** $P(X = 3)$

2 If $X \sim \text{Po}(1.8)$ find **a** $P(X = 3)$ **b** $P(X = 2)$

3 If $X \sim \text{Po}(5.3)$ find **a** $P(X = 3)$ **b** $P(X = 7)$

4 If $X \sim \text{Po}(0.4)$ find **a** $P(X = 0)$ **b** $P(X = 1)$

5 If $X \sim \text{Po}(2.15)$ find **a** $P(X = 2)$ **b** $P(X = 4)$

6 If $X \sim \text{Po}(3.2)$ find **a** $P(X = 2)$ **b** $P(X \leqslant 2)$

7 The number of telephone calls arriving at an office switchboard in a 5-minute period may be modelled by a Poisson distribution with parameter 3.2. Find the probability that in a 5-minute period

a exactly 2 calls are received **b** more than 2 calls are received.

8 The number of accidents which occur on a particular stretch of road in a day may be modelled by a Poisson distribution with parameter 1.3. Find the probability that on a particular day

a exactly 2 accidents occur **b** fewer than 2 accidents occur.

2.2 The role of the parameter of the Poisson distribution

The mean number of events in an interval of time or space is proportional to the size of the interval.

This is true as long as the average rate of occurrences remains constant.

In Example 2 in Section 2.1 the number of cars passing a point on a road during a 5-minute period was modelled by the Poisson distribution with parameter 4.

The number of cars passing that point in a 20-minute period may be modelled by the Poisson distribution with parameter 16, and in a one-minute period it may be modelled by the Poisson distribution with parameter 0.8.

If the conditions for a Poisson distribution are satisfied in a given period, they are also satisfied for periods of different length.

The average rate must be constant throughout.

EXAMPLE 1

The number of accidents in a week on a stretch of road is known to follow a Poisson distribution with parameter 2.1. Find the probability that

a in a given week there is 1 accident
b in a two-week period there are 2 accidents
c there is 1 accident in each of two successive weeks.

---

a In one week, the number of accidents follows a Po(2.1) distribution, so
the probability of 1 accident $= \dfrac{e^{-2.1}2.1^1}{1!} = 0.257\,(3\text{ s.f.})$

b In two weeks, the number of accidents follows a Po(4.2) distribution,

so the probability of 2 accidents $= \dfrac{e^{-4.2}4.2^2}{2!} = 0.132\,(3\text{ s.f.})$

c This cannot be done directly as a Poisson distribution since it concerns what happens in each of two time periods, but these are the outcomes considered in part **a**.

So the probability this happens in two successive weeks is

$$\left(\frac{e^{-2.1}2.1^1}{1!}\right)^2 = 0.0661\,(3\text{ s.f.})$$

Compare your answers to parts **b** and **c**. The probability in part **b** is higher because it includes one accident occuring in each week and also the cases where two accidents occur in one week and no accidents occur in the other week.

EXAMPLE 2

The number of flaws in a metre length of dress material is known to follow a Poisson distribution with parameter 0.4. Find the probabilities that

a there are no flaws in a 1-metre length
b there is 1 flaw in a 3-metre length
c there is 1 flaw in a half-metre length.

a $X \sim \text{Po}(0.4) \quad \Rightarrow \text{P}(X = 0) = \dfrac{e^{-0.4} \times 0.4^0}{0!} = 0.670\,(3\text{ s.f.})$

b $Y \sim \text{Po}(1.2) \quad \Rightarrow \text{P}(Y = 1) = \dfrac{e^{-1.2} \times 1.2^1}{1!} = 0.361\,(3\text{ s.f.})$

c $Z \sim \text{Po}(0.2) \quad \Rightarrow \text{P}(Z = 1) = \dfrac{e^{-0.2} \times 0.2^1}{1!} = 0.164\,(3\text{ s.f.})$

While all three of these calculations relate to the same basic situation, they all use different Poisson distributions. Using different variables avoids confusion.

Exercise 2.2

1 The number of telephone calls arriving at an office switchboard in a 5 minute period may be modelled by a Poisson distribution with parameter 1.4. Find the probability that in a 10 minute period

a exactly 2 calls are received

b more than 2 calls are received.

2 The number of accidents which occur on a particular stretch of road in a day may be modelled by a Poisson distribution with parameter 0.3. Find the probability that during a week (7 days)

a exactly 2 accidents occur on that stretch of road

b fewer than 2 accidents occur on that stretch of road.

3 The number of letters delivered to a house on a day may be modelled by a Poisson distribution with parameter 0.8.

a Find the probability that there are 2 letters delivered on a particular day.

b The home owner is away for 3 days. Find the probability that there will be more than 2 letters waiting for him when he gets back.

4 The number of errors on a page of a booklet can be modelled by a Poisson distribution with parameter 0.2.

a Find the probability that there is exactly 1 error on a given page.

b A section of the booklet has 7 pages. Find the probability that there are no more than 2 errors in the section.

c The booklet has 25 pages altogether. Find the probability that the booklet contains exactly 6 errors altogether.

2.3 The recurrence relation for the Poisson distribution

You can calculate probabilities for a Poisson distribution in sequence using a **recurrence relation**.

EXAMPLE 1

If $X \sim \text{Po}(\lambda)$

a write down the probability that

i $X = 3$ **ii** $X = 4$

b express $P(X = 4)$ in terms of $P(X = 3)$.

---

a **i** $P(X = 3) = \dfrac{e^{-\lambda} \times \lambda^3}{3!}$

ii $P(X = 4) = \dfrac{e^{-\lambda} \times \lambda^4}{4!}$

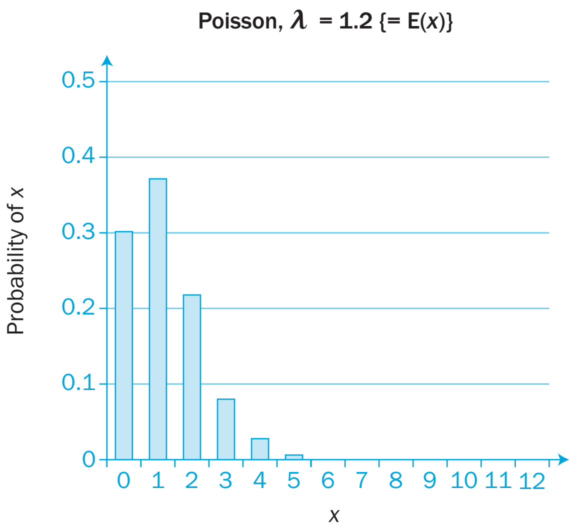
b $\dfrac{e^{-\lambda} \times \lambda^4}{4!} = \left(\dfrac{e^{-\lambda} \times \lambda^3}{3!}\right) \times \dfrac{\lambda}{4}$

$4! = 4 \times (3 \times 2 \times 1) = 4 \times 3!$

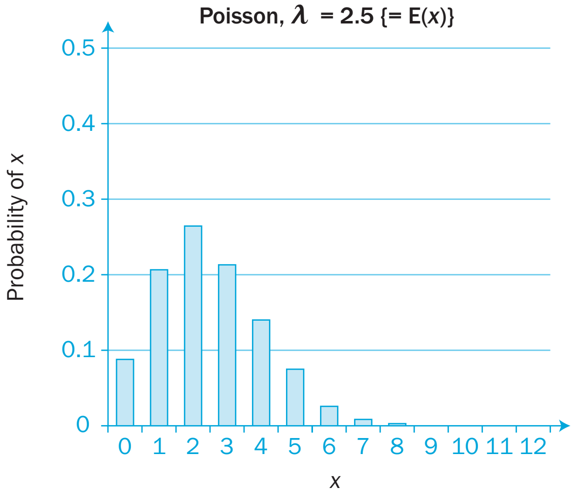
so $P(X = 4) = P(X = 3) \times \dfrac{\lambda}{4}$

The general recurrence relation for a Poisson distribution is
$P(X = k + 1) = \dfrac{\lambda}{k+1} \times P(X = k)$

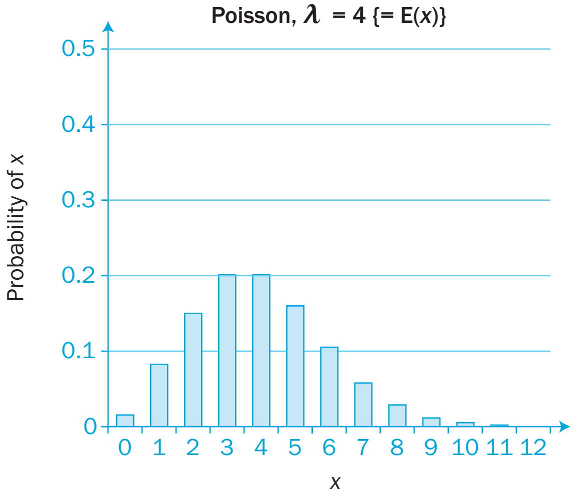
The following graphs show the probability distributions for different values of λ and what effect this has on the shape of a particular Poisson distribution.
All Poisson variables have an outcome space which is all of the non-negative integers.



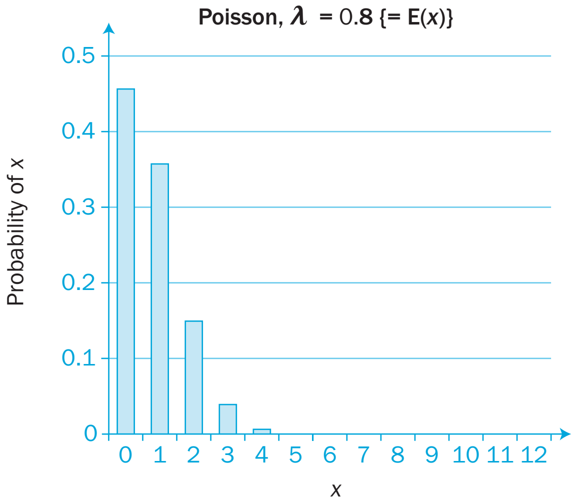
When λ is relatively low, the probabilities tail off very quickly. In this example the mode is 1.



Here λ is larger than in the previous graph, and the peak has moved to the right.
More values of X have a noticeable probability, so the highest individual probability is not as large as it was in the previous graph and the distribution is more spread out.



What happens when λ is an integer?
Generally, the mode of the Poisson (λ) distribution is at the integer below λ when λ is not an integer and there are two modes (at λ and $\lambda - 1$) when λ is an integer.



$\lambda < 1$ is a special case.
The mode will be 0 and the probability distribution is strictly decreasing for all values of X.

S2

The general forms for the probabilities of 0 and 1 for a Poisson distribution are

$$P(X = 0) = \frac{e^{-\lambda} \times \lambda^0}{0!} = e^{-\lambda} \quad \text{and} \quad P(X = 1) = \frac{e^{-\lambda} \times \lambda^1}{1!} = \lambda e^{-\lambda}$$

EXAMPLE 2

$X \sim \text{Po}(\lambda)$ and $P(X = 6) = 2 \times P(X = 5)$.
Find the value of λ.

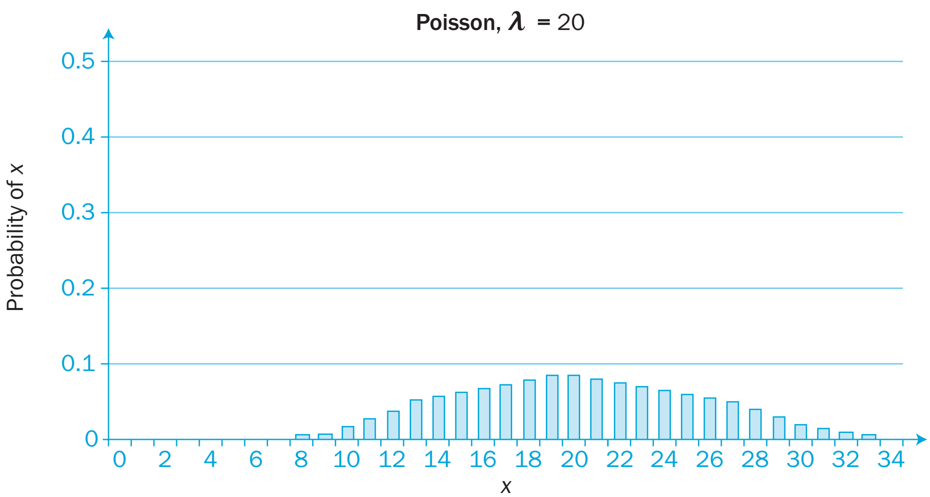
$P(X = 6) = \frac{\lambda}{6} \times P(X = 5)$ so $\frac{\lambda}{6} = 2$ and $\lambda = 12$

EXAMPLE 3

$X \sim \text{Po}(5.8)$. State the mode of X.

The mode is the integer below 5.8, i.e. the mode is 5.

This graph shows how the Poisson distribution changes when λ is large:



With $\lambda = 20$, the individual probabilities are becoming very small, and the ones which are relatively likely have moved a long way to the right of the graph – where they are still clustered around the mean (now 20).

S2

Exercise 2.3

1 $X \sim \text{Po}(2.5)$

a Write down an expression for $P(X = 4)$ in terms of $P(X = 3)$.

b If $P(X = 3) = 0.214$, calculate the value of your expression in part **a**.

c Calculate $P(X = 4)$ directly and check that it is the same as your answer to **b**.

d What is the mode of X?

2 $X \sim \text{Po}(4)$

a Write down an expression for $P(X = 4)$ in terms of $P(X = 3)$.

b Explain why X has two modes at 3 and 4.

3 $X \sim \text{Po}(\lambda)$ and $P(X = 4) = 1.2 \times P(X = 3)$

a Find the value of λ.

b What is the mode of X?

4 $X \sim \text{Po}(6.5)$

a Write down an expression for $P(X = 7)$ in terms of $P(X = 6)$.

b Explain why X has its mode at 6.

5 $X \sim \text{Po}(\lambda)$ and $P(X = 10) = 0.9 \times P(X = 87)$

a Find the value of λ.

b What is the mode of X?

Use of Poisson probability tables

The cumulative probabilities for the Poisson distributions up to $\lambda = 10.0$ in increments of 0.5 are given in the Edexcel formulae tables. (These tables are reproduced at the back of this book.) If λ is greater than 10, or not one of the values in the table, you will need to use the formula to calculate probabilities.

The Poisson distribution has only one parameter, λ, so the tables are simpler than for the binomial distribution.

POISSON CUMULATIVE DISTRIBUTION FUNCTION

The tabulated value is $P(X \leqslant x)$, where X has a Poisson distribution with parameter λ.

$\lambda =$	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
$x = 0$	0.6065	0.3679	0.2231	0.1353	0.0821	0.0498	0.0302	0.0183	0.0111	0.0067
1	0.9098	0.7358	0.5578	0.4060	0.2873	0.1991	0.1359	0.0916	0.0611	0.0404
2	0.9856	0.9197	0.8088	0.6767	0.5438	0.4232	0.3208	0.2381	0.1736	0.1247
3	0.9982	0.9810	0.9344	0.8571	0.7576	0.6472	0.5366	0.4335	0.3423	0.2650
4	0.9998	0.9963	0.9814	0.9473	0.8912	0.8153	0.7254	0.6288	0.5321	0.4405
5	1.0000	0.9994	0.9955	0.9834	0.9580	0.9161	0.8576	0.7851	0.7029	0.6160
6	1.0000	0.9999	0.9991	0.9955	0.9858	0.9665	0.9347	0.8893	0.8311	0.7622
7	1.0000	1.0000	0.9998	0.9989	0.9958	0.9881	0.9733	0.9489	0.9134	0.8666
8	1.0000	1.0000	1.0000	0.9998	0.9989	0.9962	0.9901	0.9786	0.9597	0.9319
9	1.0000	1.0000	1.0000	1.0000	0.9997	0.9989	0.9967	0.9919	0.9829	0.9682
10	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9990	0.9972	0.9933	0.9863
11	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9991	0.9976	0.9945

This is the cumulative probability distribution for the $X \sim Po(2.0)$ distribution. The number highlighted in that column has $x = 4$, and gives the probability $P(X \leqslant 4)$.

To find the probability that $X = 4$, you can use
$P(X = 4) = P(X \leqslant 4) - P(X \leqslant 3)$
which in this example would be
$0.9473 - 0.8571 = 0.0902$

Take care with the accuracy of this probability.

As with the binomial distribution, when you work with probabilities for $<$, $>$ and $\geqslant$ you need to be careful about using the correct cumulative probability.

EXAMPLE 1

If $X \sim \text{Po}(8.5)$ find
a $P(X < 7)$ **b** $P(X \geqslant 9)$ **c** $P(X > 4)$ **d** $P(4 < X < 9)$

Use the $X \sim \text{Po}(8.5)$ tables:

a $P(X < 7) = P(X \leqslant 6) = 0.2562$

b $P(X \geqslant 9) = 1 - P(X \leqslant 8) = 1 - 0.5231 = 0.4769$

c $P(X > 4) = 1 - P(X \leqslant 4) = 1 - 0.0744 = 0.9256$

d $P(4 < X < 9) = P(X < 9) - P(X \leqslant 4) = 0.5231 - 0.0301 = 0.493$

Exercise 2.4

Use the Poisson tables for this exercise.

1 If $X \sim \text{Po}(2.5)$ find **a** $P(X \leqslant 3)$ **b** $P(X < 7)$ **c** $P(3 < X < 7)$

2 If $X \sim \text{Po}(6.0)$ find **a** $P(X < 5)$ **b** $P(X > 7)$ **c** $P(5 < X \leqslant 7)$

3 If $X \sim \text{Po}(10)$ find **a** $P(X \geqslant 14)$ **b** $P(X \leqslant 18)$ **c** $P(14 < X < 18)$

4 If $X \sim \text{Po}(6.5)$ find **a** $P(X > 3)$ **b** $P(X \leqslant 6)$ **c** $P(3 \leqslant X \leqslant 6)$

5 If $X \sim \text{Po}(4.5)$ find **a** $P(X \geqslant 4)$ **b** $P(X \leqslant 4)$ **c** $P(4 \leqslant X \leqslant 4)$

6 The number of telephone calls arriving at an office switchboard in a 5 minute period may be modelled by a Poisson distribution with parameter 2.5. Find the probability that
a in a 5-minute period exactly 4 calls are received
b fewer than 6 calls are received in a quarter of an hour.

7 The number of accidents which occur on a particular stretch of road in a day may be modelled by a Poisson distribution with parameter 0.5. Find the probability that during a week (7 days)
a exactly 5 accidents occur on that stretch of road
b not more than 5 accidents occur on that stretch of road.

8 The number of letters delivered to a house on a day may be modelled by a Poisson distribution with parameter 2.5.
a Find the probability that there are 2 letters delivered on a particular day.
b The home owner is away for 3 days. Find the probability that there will be more than 8 letters waiting for him when he gets back.

9 The number of errors on a page of a booklet can be modelled by a Poisson distribution with parameter 0.2. The booklet has 40 pages altogether. Find the probability that the booklet contains at least 10 errors.

2.5 Mean and variance of the Poisson distribution

If $X \sim \text{Po}(\lambda)$, then $\text{E}(X) = \lambda$; $\text{Var}(X) = \lambda \Rightarrow$ standard deviation, $\sigma = \sqrt{\lambda}$

A special property of the Poisson distribution is that the mean and variance are always equal.

EXAMPLE 1

The number of calls arriving at a switchboard in a 10-minute period can be modelled by a Poisson distribution with parameter 3.5. Give the mean and variance of the number of calls which arrive in

a 10 minutes **b** an hour **c** 5 minutes

---

a $\lambda = 3.5$ so the mean and variance will both be 3.5.

b $\lambda = 21$ $(= 3.5 \times 6)$ so the mean and variance will both be 21.

c $\lambda = 1.75$ $(= 3.5 \div 2)$ so the mean and variance will both be 1.75.

S2

EXAMPLE 2

A dual carriageway has one lane blocked off due to roadworks. The number of cars passing a point in a road in a number of one-minute intervals is summarised in the table.

Number of cars	0	1	2	3	4	5	6
Frequency	3	4	4	25	30	3	1

a Calculate the mean and variance of the number of cars passing in one minute intervals.

b Is the Poisson distribution likely to be an adequate model for the distribution of the number of cars passing in one-minute intervals?

---

a $\sum f = 70$, $\sum xf = 228$ $\sum x^2 f = 836$, so $\bar{x} = \frac{228}{70} = 3.26$ (3 s.f.)

and $\text{Var}(X) = \frac{\sum x^2 f}{\sum f} - \bar{x}^2 = \frac{836}{70} - \left(\frac{228}{70}\right)^2 = 1.33$ (3 s.f.)

b The mean and variance are not numerically close so it is unlikely the Poisson will be an adequate model (with only one lane open for traffic, overtaking cannot happen on this stretch of the road and the numbers of cars will be much more consistent than normal – hence the variance is much lower than would be expected if the Poisson model did apply).

Derivation of mean and variance of the Poisson distribution

Let $X \sim \text{Po}(\lambda)$

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

$$E(X) = \sum_{k=0}^{\infty} k \times \frac{e^{-\lambda}\lambda^k}{k!} = \lambda \times \sum_{k=1}^{\infty} \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}$$

$$= \lambda$$

As in the binomial case, cancel *k*, after discarding the zero case.

$$E(X^2) = \sum_{k=0}^{\infty} k^2 \times \frac{e^{-\lambda}\lambda^k}{k!} = \lambda \times \sum_{k=1}^{\infty} k \times \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}$$

$$= \lambda \times \sum_{k=1}^{\infty} (k-1+1) \times \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}$$

$$= \lambda^2 \times \sum_{k=2}^{\infty} \frac{e^{-\lambda}\lambda^{k-2}}{(k-2)!} + \lambda \times \sum_{k=1}^{\infty} \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}$$

$$= \lambda^2 + \lambda$$

Then $\text{Var}(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$

You are not expected to learn this derivation.

S2

Exercise 2.5

1 If $X \sim \text{Po}(3.2)$ find **a** $E(X)$ **b** $\text{Var}(X)$

2 If $X \sim \text{Po}(49)$ find the mean and standard deviation of X.

3 $X \sim \text{Po}(3.5)$. Find

a the mean and standard deviation of X

b $P(X > \mu)$ where $\mu = E(X)$

c $P(X > \mu + 2\sigma)$, where σ is the standard deviation of X

d $P(X < \mu - 2\sigma)$.

4 X is the number of telephone calls arriving at an office switchboard in a ten-minute period. X may be modelled by a Poisson distribution with parameter 6. Find

a the mean and standard deviation of X

b $P(X > \mu)$, where $\mu = E(X)$

c $P(X > \mu + 2\sigma)$, where σ is the standard deviation of X

d $P(X < \mu - 2\sigma)$.

5 What can you gather from your answers to part **d** of questions **3** and **4**?

2.6 Modelling with the Poisson distribution

The Poisson distribution describes the number of occurrences in a fixed period of time or space if the events occur independently of one another, at random and at a constant average rate.

Standard examples of Poisson processes in real life include radioactive emissions, traffic passing a fixed point, telephone calls or letters arriving, and accidents occurring.

EXAMPLE 1

The maternity ward of a hospital wanted to work out how many births would be expected during a night.

The hospital had 3000 deliveries each year so, if these happened randomly around the clock, 1000 deliveries would be expected between the hours of midnight and 8.00 a.m. This is the time when many staff are off duty and it is important to ensure that there will be enough people to cope with the workload on any particular night.

The average number of deliveries per night is $\frac{1000}{365}$, which is 2.74.

From this average rate you can calculate the probability of delivering 0, 1, 2, etc. babies each night using the Poisson distribution. Some probabilities are:

$P(0) = 2.74^0 \frac{e^{-2.74}}{0!} = 0.065$ $\quad P(2) = 2.74^2 \frac{e^{-2.74}}{2!} = 0.242$

$P(1) = 2.74^1 \frac{e^{-2.74}}{1!} = 0.177$ $\quad P(3) = 2.74^3 \frac{e^{-2.74}}{3!} = 0.221$

a On how many nights in the year would 5 or fewer deliveries be expected?
b Over the course of one year, what is the greatest number of deliveries expected in any night?
c Why might the pattern of deliveries *not* follow a Poisson distribution?

---

a Let X = number of deliveries

$365 \times P(X \leqslant 5) = 343$

b 8
This is the largest value for which the probability is greater than $\frac{1}{365}$.

c If deliveries were not random throughout the 24 hours.
E.g. if a lot of women had elective caesareans done during the day.

Real-life example.
By kind permission of the Mathematics Association.

S2

In the real-life example described in Example 1, deliveries in fact followed the Poisson distribution very closely, and the hospital was able to predict the workload accurately.

Remember the conditions for the Poisson distribution:

1 events occur at random
2 events occur independently of one another
3 the average rate of occurrences remains constant
4 there is zero probability of simultaneous occurrences.

Be careful:
Some change in the underlying conditions may alter the nature of the distribution.
E.g. Traffic observed close to a junction, or where there are lane restrictions and traffic is funnelled into a queue travelling at constant speed.

The underlying conditions may be distorted by interference from other effects.
E.g. If a birthday or Christmas occurs during the period considered then the Poisson conditions would not be reasonable for the arrival of letters.

Randomness or independence may be lost because of a difference in the average rate of occurrences.
E.g. The rate of accidents occurring would be expected to vary somewhat as road conditions vary.

S2

EXAMPLE 2

The number of cyclists passing a village post office during the day can be modelled as a Poisson random variable.
On average two cyclists pass by in an hour.
What is the probability that

a between 10.00 a.m. and 11.00 a.m.

  i no cyclist passes

  ii more than 3 cyclists pass

b exactly one cyclist passes while the shop-keeper is on a 20-minute tea-break

c more than 3 cyclists pass in an hour exactly once in a six-hour period?

---

a In an hour (parts **i** and **ii**) $\lambda = 2$.

  i $P(X = 0) = 0.1353$ (either by using the tables, or by calculating e^{-2}).

  ii $P(X > 3) = 1 - P(X \leqslant 3)$
  $= 1 - 0.8571$
  $= 0.1429$ (from tables)

Example 2 is continued on the next page.

EXAMPLE 2 (CONT.)

b In a 20-minute period $\left(\frac{1}{3} \text{ of an hour}\right)$,

the mean number of cyclists will be $2 \times \frac{1}{3} = \frac{2}{3}$,
which is not in the tables.

$$\text{P(exactly 1)} = \frac{e^{-\frac{2}{3}}\left(\frac{2}{3}\right)^1}{1!} = 0.342 \text{ (3 s.f.)}$$

c The situation is that of a binomial distribution.
There are six 'trials', the number of cyclists in each hour is independent of the other periods, and the probability of more than 3 in an hour remains the same for all the six hour-long periods. Therefore if Y is the number of times that more than 3 cyclists pass by in an hour exactly once in a six-hour period $Y \sim \text{B}(6, 0.1429)$
$\text{P}(Y = 1) = 0.3966$

This cannot be treated as a single Poisson distribution with parameter 12 since it specifies a particular event to be considered in each one-hour time period separately.

Using the probability calculated in part **a ii**.

Exercise 2.6

S2

1 For the following random variables state whether they can be modelled by a Poisson distribution.
If they can, give the value of the parameter λ; if they cannot then explain why.

a The average number of cars per minute passing a point on a road is 12.
The traffic is flowing freely.
X = number of cars which pass in a 15-second period.

b The average number of cars per minute passing a point on a road is 14.
There are roadworks blocking one lane of the road.
X = number of cars which pass in a 30-second period.

c Amelie normally gets letters at an average rate of 1.5 per day.
X = number of letters Amelie gets on 22 December.

d A petrol station which stays open all the time gets an average of 832 customers in a 24-hour time period.
X = number of customers in a quarter of an hour at the petrol station.

e An Accident and Emergency department in a hospital treats 32 patients an hour on average.
X = number of patients treated between 1 a.m. and 2 a.m. on a Sunday morning.

You are *not* expected to do any calculations.

2 For the following situations state what assumptions are needed if a Poisson distribution is to be used to model them, and give the value of λ that would be used.

a On average, defects in a roll of cloth occur at a rate of 0.2 per metre.
How many defects are there in a roll which is 8 m long?

b On average, misprints on a page in a magazine occur once in 2 paragraphs.
How many errors are there in a page with 8 paragraphs?

c A small shop averages 8 customers per hour.
How many customers does it have in 20 minutes?

3 An explorer thinks that the number of mosquito bites he gets when he is in the jungle will follow a Poisson distribution.

The explorer records the number of mosquito bites he gets in the jungle during a number of hour-long periods.
The results are summarised in the table.

Number of bites	0	1	2	3	4	5	6	$\geqslant 7$
Frequency	3	7	9	6	6	3	1	0

a Calculate the mean and variance of the number of bites the explorer gets in an hour in the jungle.

b Do you think the Poisson distribution is a good model for the number of bites the explorer gets in an hour in the jungle?

4 The number of emails Serena gets can be modelled by a Poisson distribution with a mean rate of 1.5 per hour.
What is the probability that

a i Serena gets no emails between 4 p.m. and 5 p.m.?
ii Serena gets more than 2 emails between 4 p.m. and 5 p.m.?
iii Serena gets one email between 6 p.m. and 6.20 p.m.?
iv Serena gets more than 2 emails in an hour exactly twice in a five-hour period.?

b Would it be sensible to use the Poisson distribution to find the probability that Serena gets no emails between 4 a.m. and 5 a.m?
Explain why.

S2

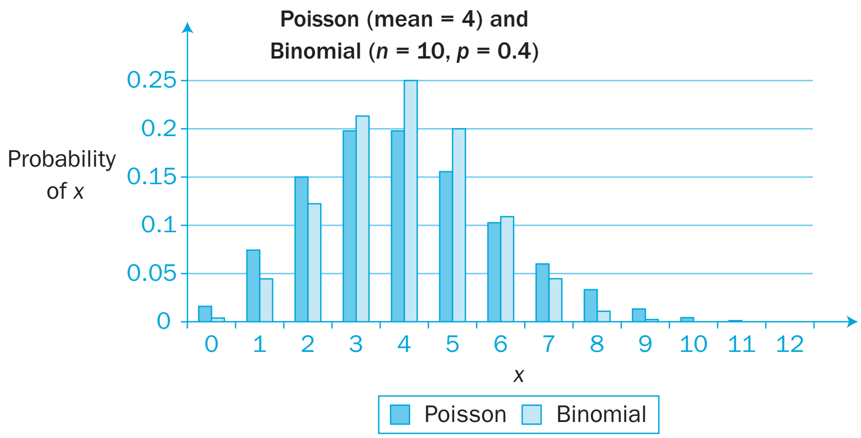
2.7 Poisson approximation to the binomial distribution

If $X \sim B(n, p)$ with n large and p close to 0 then
$X \sim$ *approximately* $Po(\lambda)$ with $\lambda = np$

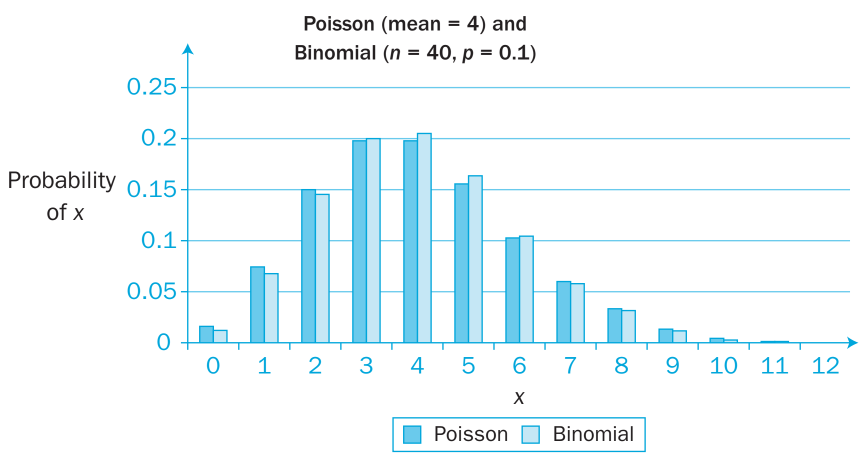
If p is close to 0,
then $q = 1 - p$ will be close to 1
and npq will be close to np.

np and npq are the mean and variance of the binomial distribution, and the Poisson has the property that the mean and variance are equal.

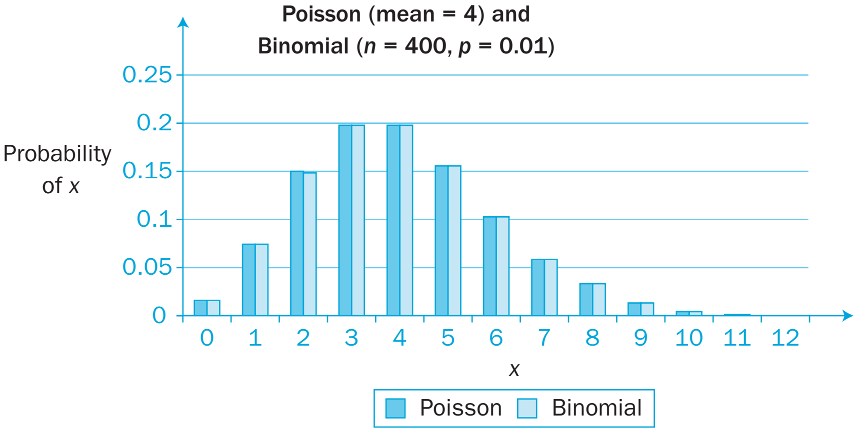
Here are four examples where the binomial and Poisson distributions have the same mean:

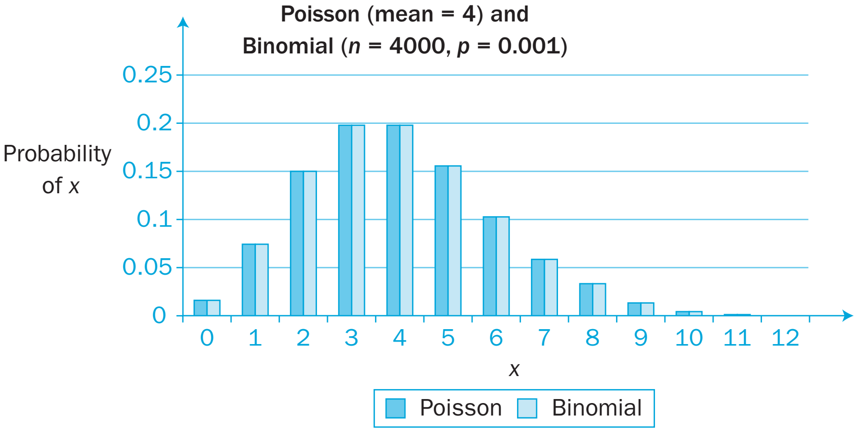


Here the mean of the binomial is 4 and the variance is 2.4.
The two sets of probabilities are not particularly similar.



In this case, the variance of the binomial is 3.6. The variance of the Poisson is 4.
The agreement between the two sets of probabilities is quite strong.





In these two examples, while you cannot see any difference on this scale graphically, there are differences between the binomial and the Poisson in both cases. The differences in the second graph are much smaller than the differences in the first graph.

The use of the Poisson distribution as an approximation to the binomial distribution improves as n increases and as p gets smaller.

The Poisson outcome space has no upper limit whereas the binomial is bounded by the value of n. However, when n is large and p is small, the probabilities of high values of x are very small so this is not a problem.

S2

EXAMPLE 1

The probability that a component coming off a production line is faulty is 0.01.

a If a sample of size 5 is taken, find the probability that one of the components is faulty.

b What is the probability that a batch of 250 of these components has more than 3 faulty components in it?

---

a If X is the number of faulty components in the sample then $X \sim B(5, 0.01)$ and

$$P(X = 1) = 5 \times 0.01 \times 0.99^4$$
$$= 0.0480 \text{ (3 s.f.)}$$

b If Y is the number of faulty components in the batch then

$$X \sim B(250, 0.01) \overset{approx.}{\sim} Po(2.5)$$

and $P(Y > 3) = 1 - P(Y \leqslant 3)$
$= 1 - 0.758$
$= 0.242$

$P(X > 3) = 0.242$ if $B(250, 0.01)$ is used.

If you are working in a situation where p is close to 1, you can choose to count failures instead of successes and still construct an appropriate Poisson approximation.

S2

Exercise 2.7

1 The proportion of defective pipes coming off a production line is 0.05. A sample of 40 pipes are examined.

a Using the tables of an exact binomial distribution calculate the probabilities that there are

i 0 **ii** 1

iii 2 **iv** more than 2

defectives in the sample.

b Using an appropriate approximate distribution calculate the probabilities that there are

i 0 **ii** 1

iii 2 **iv** more than 2

defectives in the sample.

2 a State the conditions under which a Poisson distribution may be used to approximate a binomial distribution.

b 5% of the times a faulty ATM asks for a personal identification number (PIN) it does not register the number entered correctly. If a customer enters their PIN correctly each time, what is the probability that the ATM will not register it correctly in 3 attempts?

c Over a period of time, 90 attempts are made to enter a PIN. If all of the customers enter their PIN correctly, what is the probability that fewer than 3 of the attempts are not registered correctly?

3 In a small town, the football team claims that 95% of the people in town support them. If a survey of 200 randomly chosen people asks whether they support the football team, find the probability that more than 195 people say they do.

4 A rare but harmless medical condition affects 1 in 200 people.

a In a cinema in which 130 people are watching a film, what is the probability that exactly one person has the condition?

b At a concert where the audience is 1800, use an appropriate approximate distribution to find the probability that there are fewer than 5 people with the condition.

5 The Nutty Fruitcase Party claim that 1 in 250 people support their policy to distribute free fruit and nut chocolate bars to children taking examinations.

a In an opinion poll which asks 1000 voters about a range of policies put forward by different parties what is the probability that

i no one will support the Nutty Fruitcase Party policy
ii at least 10 people will support the policy.

b If the opinion poll had 7 people supporting the policy, does this mean that the Nutty Fruitcase Party have underestimated the support there is for this policy?

Review 2

1 Customers enter a shop independently of one another, and at random intervals of time, at an average rate of 5 per hour during the time that the shop is open.

a If X is the number of customers entering the shop in a one-hour period, write down the distribution that you would use to model X.

b Use this distribution to calculate the probability that

i there are 4 customers in one particular hour

ii there are more than 6 customers in a given hour

iii there are fewer than 9 customers in a two-hour period

iv exactly one customer enters in each of four 15-minute periods.

c Compare the probabilities obtained in parts **i** and **iv** of **b**.

2 Lemons are packed in bags containing 5 each. It is found that, on average, 6% of the lemons are too sour to use.

a Find the probability that a bag contains

i one unusable lemon ii more than one unusable lemon.

b A box containing 15 of these bags is opened and inspected. Identify and use a suitable approximation to find the probability that there are no more than 3 unusable lemons.

3 A shop sells a particular make of DVD player.

a Assuming that the weekly demand for the DVD player is a Poisson variable with mean 5, find the probability that the shop sells

i at least 4 in a week

ii at most 7 in a week

iii more than 12 in a fortnight (2 weeks).

b Stocks are brought in only at the beginning of each fortnight. Find the minimum number that should be in stock at the beginning of a fortnight so that the shop can be at least 95% sure of being able to meet all demands during the fortnight.

4 A car hire firm finds that the daily demand for its cars follows a Poisson distribution with mean 4.5.

a What is the probability that on a particular day the demand is

i 2 or fewer ii between 4 and 9 (inclusive) iii zero?

b What is the probability that 8 consecutive days will include two or more on which the demand is zero?

c Suggest reasons why the daily demand for car hire may not follow a Poisson distribution.

5 Serious accidents in a certain type of manufacturing industry can be adequately modelled by the Poisson distribution with a mean rate of 1.4 per week.

a What is the probability that there are no serious accidents in a particular week?

b What is the probability that there are at least three serious accidents in a three-week period?

c What is the probability that in a four-week period there is exactly one week in which there are serious accidents?

6 A firm is proud of their production statistics, which show that only 0.15% of their components are faulty. The components are packed in boxes of 500.

a Write down the appropriate exact distribution to model the number of faulty components in a box chosen at random.

i Give two conditions which need to be satisfied for this distribution to be a suitable model.

ii Is it reasonable to assume that these conditions would be satisfied?

b Write down the distribution you would use to calculate the probability that a box contains more than one faulty component.

c Calculate this probability.

7 A computer repair company uses a particular spare part at a rate of 3 per week. Assuming that requests for this spare part occur at random, find the probability that

a exactly 4 are used in a particular week

b at least 10 are used in a two-week period

c exactly 4 are used in each of three consecutive weeks.

The manager decides to replenish the stock of this spare part to a constant level n at the start of each week.

d Find the value of n such that, on average, the stock will be insufficient no more than once in a 52 week year.

8 An Internet site is visited on average 5 times an hour during the period between 10 a.m. and 4 p.m., and the visits occur at random intervals of time, independently of one another. What distribution is appropriate to model the numbers of visits recorded between 11 a.m. and 1 p.m? Calculate the probability that there are

a fewer than 4 visits between 11 a.m. and 1 p.m.

b exactly 2 visits between 11 a.m. and 11.30 a.m.

9 a Serious accidents in a large factory can be adequately modelled by the Poisson distribution with a mean rate of 1.3 per month.

i What is the probability that there are no serious accidents in a particular month?

ii What is the probability that in a four-month period there is at most one month during which there are serious accidents?

b i Give *two* assumptions of the Poisson distribution.

ii Comment on whether you think the Poisson would be the exact distribution for the number of serious accidents in a month.

10 Customers enter a large cafeteria either alone or in groups.

a The number of customers entering alone between 10.00 a.m. and 11.00 a.m. may be modelled by a Poisson distribution with a mean of 0.6 per minute.

Find the probability that, during a particular minute between 10.00 a.m. and 11.00 a.m., the number of customers entering the cafeteria alone is

i 2 or fewer ii exactly 3.

b The number of groups of customers entering the cafeteria between 10.00 a.m. and 11.00 a.m. may be modelled by a Poisson distribution with mean of 0.3 per minute. Find the probability that

i during a particular minute between 10.00 a.m. and 11.00 a.m., more than one group of customers enters the shop

ii 4 or more groups of customers enter the cafeteria between 10.30 a.m. and 10.45 a.m.

c Explain whether the Poisson distribution is likely to provide a suitable model for the *number of customers* entering the cafeteria in groups during each minute between 10.00 a.m. and 11.00 a.m.

d The cafeteria is open from 8.00 a.m. until 6.00 p.m. Explain whether the Poisson distribution is likely to provide a suitable model for the number of customers entering the cafeteria alone during each minute when the cafeteria is open.

11 Rachelle sells a magazine which is produced to raise money for homeless people. The probability of making a sale is 0.05 for each person she approaches.

a Given that she approaches 30 people, find the probability that she will make
  i 2 or fewer sales
  ii more than 4 sales.

b Find the probability that she makes 2 sales given that she approaches 24 people.

c State *one* assumption you have made.

12 The average number of calls to a telephone exchange during the half hour from 10 a.m. to 10.30 a.m. on weekdays is six. Find the probability that

You may assume that the Poisson model is adequate.

a on a given weekday between these times, the exchange will receive exactly four calls.

b on any given weekday, the exchange will receive more than three calls between 10 a.m. and 10.10 a.m.

13 The number of vehicles arriving at a toll bridge during a 5-minute period can be modelled by a Poisson distribution with mean 2.4.

a State the value for the standard deviation of the number of vehicles arriving at the toll bridge during a 5-minute period.

b Find the probability that
  i at least 2 vehicles arrive in a 5-minute period.
  ii at least 2 vehicles arrive in each of three successive 5-minute periods.

c Show that the probability that *no* vehicles arrive in a 10 minute period is 0.0082, correct to four decimal places.

2 Exit

Summary

Refer to

- The Poisson distribution has one parameter, λ.
  - $P(X = r) = \frac{e^{-\lambda}\lambda^r}{r!}$ for $r = 0, 1, 2, 3, \ldots$ 2.1
- The Poisson distribution requires certain conditions:
  - events occur at random
  - events occur independently of one another
  - the average rate of occurrences remains constant
  - there is zero probability of simultaneous occurrences. 2.1

  If these conditions are not met completely, the Poisson distribution may still provide a reasonable model. 2.6
- The value of the parameter is proportional to the length of the interval (of time or space). 2.2
- If $X \sim \text{Po}(\lambda)$ then $P(X = k + 1) = \frac{\lambda}{k+1} \times P(X = k)$ 2.3
- The probability tables for the Poisson distribution give the cumulative probability for certain values of λ for values of λ in increments of 0.5 up to 10. 2.4
- If $X \sim \text{Po}(\lambda)$, then $E(X) = \lambda$; $\text{Var}(X) = (\sigma^2) = \lambda$ 2.5
- If $X \sim B(n, p)$ with n large and p close to 0 then $X \sim$ *approximately* $\text{Po}(\lambda)$ with $\lambda = np$ 2.7

Links

The Poisson distribution can be used to model a large number of natural and social phenomena as well as providing a good approximation to large binomial distributions under certain conditions.

The number of genetic mutations in a stretch of DNA may be modeled by the Poisson distribution.

This can help in the understanding of mutations in both plants and animals and can be applied to medical research in diseases such as cancer and Parkinson's.

S2