Modelling with the Poisson Distribution

The Poisson distribution is a probability distribution used to find the probability that an event will occur x number of times within a certain interval. It models events that occur:

- **singly** (one at a time).
- **independently** (an event occurring will not affect the probability of another event occurring).
- at a **constant average rate** (the average number of times the event occurs in an interval is fixed).

Examples where a Poisson distribution may be used include:

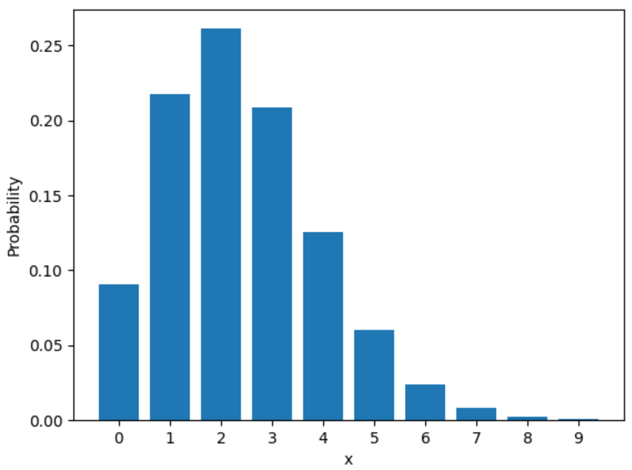
- the number of cars passing point on a motorway in 10 minutes,
- the number of worms in a $1\ \text{m}^2$ patch of grass,
- the number of faulty items produced in a factory per week,
- the number of times a radioactive substance decays in 1 minute.

If X is a discrete random variable that follows the Poisson distribution, the mathematical notation used is as follows:

$$X \sim Po(\lambda)$$

λ is the only parameter, it represents the mean number of times an event will occur in an interval of space or time. Notice that λ does not need to be an integer.

A standard Poisson distribution (for $\lambda = 2.4$) is shown below. Unlike the binomial distribution, the Poisson distribution is not symmetrical and is defined for all positive integer values of x (though as x tends to ∞, the probabilities tend to 0).



Example 1: 1.875×10^{20} electrons pass through a 1 m length of wire in one minute. If X represents the number of electrons passing the wire in one second, write down the distribution of X, stating suitable assumptions.

Calculate λ – the number of electrons passing the wire in one second.	$\lambda = \frac{1.875 \times 10^{20}}{60} = 3.125 \times 10^{18}$
Write down the distribution of X.	$X \sim Po(3.125 \times 10^{18})$
State the assumptions for a Poisson distribution **in context**.	• The electrons pass through the wire one at a time • The electrons pass through the wire independently of each other. • The electrons pass through the wire at a constant average rate.

Calculating Probabilities Using the Poisson Function or a Calculator

Poisson probabilities can be calculated using the following formula:

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

where $x \in \mathbb{Z}^+ (x = 0,1,2 \ldots)$ and $e = 2.718 \ldots$ is Euler's number.

The 'Poisson PD' and 'Poisson CD' functions on a calculator can also be used.

Example 2: The number of errors in an essay is modelled by a Poisson distribution with 3.8 errors per page on average. Find the probability that there are:

a) no errors on the next page.

b) fewer than 2 errors on the fifth page.

c) more than 7 errors on the last 2 pages.

a) Define the random variable X in context and write its distribution.	X represents the number of errors made on a page of an essay. $X \sim Po(3.8)$
Use the 'Poisson PD' function on the calculator with $x = 0$ and $\lambda = 3.8$.	$P(X = 0) = 0.0224\ (3\ s.f.)$
b) Rewrite $P(X < 2)$ as $P(X \leq 1)$, ready to be entered into the calculator. Use the 'Poisson CD' function with $x = 1$ and $\lambda = 3.8$.	$P(X < 2) = P(X \leq 1) = 0.1074$
c) The Poisson distribution is scalable so if 3.8 errors are made on one page on average, 7.6 errors will be made on two pages. Define a new random variable Y in context, and write its distribution. Rewrite $P(Y > 7)$ as $1 - P(Y \leq 7)$, ready to be entered into the calculator. Use the 'Poisson CD' function with $y = 7$ and $\lambda = 7.6$.	Y represents the number of errors made on two pages of an essay. $Y \sim Po(7.6)$ $P(Y > 7) = 1 - P(Y \leq 7)$ $= 1 - 0.5100\ldots$ $= 0.490\ (3\ s.f.)$

Example 3: The random variable X follows a Poisson distribution with mean λ. If $P(X = 3) = \frac{P(X=0)}{8} - P(X = 6)$, find the value of λ.

Rewrite probabilities using the Poisson formula.	$\frac{e^{-\lambda} \times \lambda^3}{3!} = \frac{1}{8}\left(\frac{e^{-\lambda} \times \lambda^0}{0!}\right) - \frac{e^{-\lambda} \times \lambda^6}{6!}$
Simplify the equation by dividing through by $e^{-\lambda}$ and multiplying through by 6!. Notice $e^{-\lambda} > 0$ for all λ so this is valid.	$\frac{\lambda^3}{3!} = \frac{1}{8}\left(\frac{\lambda^0}{0!}\right) - \frac{\lambda^6}{6!}$ $120\lambda^3 = 90 - \lambda^6$
Recognise that the equation is a hidden quadratic and solve for λ^3 using the quadratic solver function.	$\lambda^6 + 120\lambda^3 - 90 = 0$ $(\lambda^3)^2 + 120\lambda^3 - 90 = 0$ Polynomial solver: $\lambda^3 = 0.7454, -120.7$
Deduce the two possible values of λ, and exclude the negative value.	$\lambda = 0.907\ (3\ s.f.)$ as $-4.94 < 0$

Mean, Variance and Standard Deviation of the Poisson Distribution

If $X \sim Po(\lambda)$:

$$E(X) = \lambda$$

$$Var(X) = \sigma^2 = \lambda$$

An indication that data may be well modelled by a Poisson distribution is if $\lambda \approx E(X) \approx Var(X)$. In theory, these three values should be the same.

Example 4: A student is investigating the number of birds that land on a hedge in 5 minutes by taking 100 random samples. The results are summarised as follows:

$$\Sigma x = 2348, \Sigma x^2 = 57483$$

a) Calculate the mean and the variance of the number of birds that land on the hedge in 5 minutes.

b) Explain whether a Poisson distribution could be used to model this data, using your answers from part **a**.

c) Using a suitable value for λ, estimate the probability that 27 birds will land on the hedge in a period of 5 minutes.

a) Calculate the mean by using the formula $E(X) = \frac{\Sigma x}{n}$.	$E(X) = \frac{\Sigma x}{n} = \frac{2348}{100} = 23.48$
Calculate the variance by using the formula $Var(X) = E(X^2) - [E(X)]^2$.	$Var(X) = E(X^2) - [E(X)]^2 = \frac{57483}{100} - 23.48^2$ $= 23.52$
b) Identify that the mean and variance are approximately equal and draw a conclusion from this.	$23.48 \approx 23.52$. As the mean and the variance are approximately equal, a Poisson distribution could be used to model this data.
c) Define the random variable X and write the distribution with $\lambda = 23.48$ (it would have also been valid to let $23.38 \leq \lambda \leq 23.52$). Use the 'Poisson PD' function on a calculator with $x = 27$ and $\lambda = 23.48$ to find the required probability.	Let X represent the number of birds that land on the hedge in 5 minutes. $X \sim Po(23.48)$ $P(X = 27) = 0.0595\ (3\ s.f.)$