

## The Poisson distribution

The Poisson distribution is, like the Binomial distribution, a theoretical discrete distribution that is used to calculate the probability that a random event occurs a given number of times. Given a Poisson variable  $X$ , the values assumed by  $X$  have probabilities that can be generated by its *probability mass function*

$$P[X = x] = \frac{e^{-\mu} \mu^x}{x!} \text{ for } x = 0, 1, 2, 3, \dots$$

where  $X$  is defined as the *number of occurrences* of the variable and  $\mu$  is the mean number of occurrences over a specific *interval* (explained below). In fact,  $\mu$  is the only parameter of the Poisson distribution and also happens to be the *variance* of the distribution. We hence write

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



Unlike the case for the binomial distribution, we do not have an upper limit for the number of occurrences of a randomly occurring event (which, as you will see further, makes a lot of sense). In some books, the parameter  $\lambda$  is used instead of  $\mu$ .

The **conditions** for applying the Poisson distribution to a given problem are

1. The occurrences of the event are completely independent of one another (that is, they are completely *random*)
2. The number of occurrences in a given *space-time interval* is independent of those occurring in another interval.

An example of a randomly occurring event is “the number of earthquakes in Turkey in a one year-period”. It is seen in this example that the *time interval* has to be defined even though we know that our variable is the number of earthquakes in Turkey. This is what a ‘space-time’ interval is all about! Other variables could be

1. Number of breaks in a white line on a 1-kilometre road (*length*)
2. Number of misprints per page in a book (*area*)
3. Number of micro-organisms in 10 cm<sup>3</sup> of liquid (*volume*)
4. Number of customers arriving in a bank per minute (*time*)

Now that we understand the concept of a space-time interval, let us have a look at an elementary example.

*Example 1*

The number of emergency admissions each day to a hospital is found to have a Poisson distribution with mean 2.

- (a) Evaluate the probability that, on a particular day, there will be no emergency admissions.
- (b) At the beginning of one day the hospital has 5 beds available for emergencies. Calculate the probability that this will be an insufficient number for the day.
- (c) Calculate the probability that there will be exactly 3 admissions altogether on two consecutive days.

*Solution*

We should *always* start by defining our variable since its distribution varies according to its space-time interval. So if  $X = \text{“number of emergency admissions per day”}$ , we write  $X \sim \text{Po}(2)$ .

(a) 
$$P[X = 0] = \frac{e^{-2} 2^0}{0!} = 0.135$$

- (b) If the number of beds is insufficient, we can easily deduce that the number of emergencies exceeded 5. Hence we have to calculate  $P[X > 5]$ . Unfortunately, that would mean calculating an infinite number of probabilities since the Poisson variable has no upper limit for its number of occurrences. We revert to the theory of probability which states that the sum of all probabilities in a given experiment must be 1. It follows that

$$P[X > 5] = 1 - P[X \leq 5].$$

Now,  $P[X \leq 5] = P[X = 0] + P[X = 1] + \dots + P[X = 5]$

$$\begin{aligned} &= \frac{e^{-2} 2^0}{0!} + \frac{e^{-2} 2^1}{1!} + \frac{e^{-2} 2^2}{2!} + \frac{e^{-2} 2^3}{3!} + \frac{e^{-2} 2^4}{4!} + \frac{e^{-2} 2^5}{5!} \\ &= e^{-2} \left( 1 + 2 + 2 + \frac{4}{3} + \frac{2}{3} + \frac{4}{15} \right) = 0.983 \end{aligned}$$

Thus,  $P[X > 5] = 1 - 0.983 = 0.017$ .

- (c) This part of the question shows how *flexible* the Poisson distribution can be! The space-time time interval controls the parameter of the distribution. We have to define a new Poisson variable (let's call it  $Y$ ) defined as

$$Y = \text{“number of emergency admissions for a two-day period”}$$

so that  $Y \sim \text{Po}(4)$ .

$$P[Y = 3] = \frac{e^{-4} 4^3}{3!} = 0.195$$