The Poisson Distribution

October 20, 2010

In many practical situations we are interested in measuring how many times a certain event occurs in a specific time interval or in a specific length or area. For instance:

1. the number of phone calls received at an exchange or call centre in an hour;
2. the number of customers arriving at a toll booth per day;
3. the number of flaws on a length of cable;
4. the number of cars passing using a stretch of road during a day.

Poisson Distribution

In many practical situations we are interested in measuring how many times a certain event occurs in a specific time interval or in a specific length or area. For instance:

1. the number of phone calls received at an exchange or call centre in an hour;
2. the number of customers arriving at a toll booth per day;
3. the number of flaws on a length of cable;
4. the number of cars passing using a stretch of road during a day.

The Poisson distribution plays a key role in modelling such problems.

Suppose we are given an interval (this could be time, length, area or volume) and we are interested in the number of "successes" in that interval.
Assume that the interval can be divided into very small subintervals such that:

1. the probability of more than one success in any subinterval is zero;

Poisson Distribution

Suppose we are given an interval (this could be time, length, area or volume) and we are interested in the number of "successes" in that interval.
Assume that the interval can be divided into very small subintervals such that:

1. the probability of more than one success in any subinterval is zero;
2. the probability of one success in a subinterval is constant for all subintervals and is proportional to its length;

Poisson Distribution

Suppose we are given an interval (this could be time, length, area or volume) and we are interested in the number of "successes" in that interval.
Assume that the interval can be divided into very small subintervals such that:

1. the probability of more than one success in any subinterval is zero;
2. the probability of one success in a subinterval is constant for all subintervals and is proportional to its length;
3. subintervals are independent of each other.

Poisson Distribution

We assume the following.

- The random variable X denotes the number of successes in the whole interval.
- λ is the mean number of successes in the interval.

Poisson Distribution

We assume the following.

- The random variable X denotes the number of successes in the whole interval.
- λ is the mean number of successes in the interval.

X has a Poisson Distribution with parameter λ and

$$P(X = k) = f(k) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots.$$

Example

The number of flaws in a fibre optic cable follows a Poisson distribution. The average number of flaws in 50m of cable is 1.2.

(i) What is the probability of exactly three flaws in 150m of cable?
(ii) What is the probability of at least two flaws in 100m of cable?
(iii) What is the probability of exactly one flaw in the first 50m of cable and exactly one flaw in the second 50m of cable?

Example

(i) Mean number of flaws in 150m of cable is 3.6. So the probability of exactly three flaws in 150m of cable is

$$\frac{e^{-3.6}(3.6)^3}{3!} = 0.212.$$

(ii) Mean number of flaws in 100m of cable is 2.4. Let X be the number of flaws in 100m of cable.

$$\begin{aligned} P(X \geq 2) &= 1 - (P(X=0) + P(X=1)) \\ &= 1 - (\frac{e^{-2.4}(2.4)^0}{0!} + \frac{e^{-2.4}(2.4)^1}{1!}) \\ &= 1 - (0.091 + 0.218) = 0.691. \end{aligned}$$

Example

(iii) Now let X denote the number of flaws in a 50m section of cable. Then we know that

$$P(X = 1) = \frac{e^{-1.2}(1.2)^1}{1!} = 0.361.$$

As X follows a Poisson distribution, the occurrence of flaws in the first and second 50m of cable are independent. Thus the probability of exactly one flaw in the first 50m and exactly one flaw in the second 50m is

$$(0.361)(0.361) = 0.13.$$

Example

The number of visitors to a webserver per minute follows a Poisson distribution. If the average number of visitors per minute is 4, what is the probability that:

(i) There are two or fewer visitors in one minute?;

(ii) There are exactly two visitors in 30 seconds?.

Example

The number of visitors to a webserver per minute follows a Poisson distribution. If the average number of visitors per minute is 4, what is the probability that:

(i) There are two or fewer visitors in one minute?;

(ii) There are exactly two visitors in 30 seconds?.

(i) For part (i), we need the average number of visitors in a minute. In this case the parameter $\lambda = 4$.
We wish to calculate

$$P(X = 0) + P(X = 1) + P(X = 2).$$

Example

$$
\begin{aligned}
P(X=0) &= \frac{e^{-4}4^0}{0!} = e^{-4} \\
P(X=1) &= \frac{e^{-4}4^1}{1!} = 4e^{-4} \\
P(X=2) &= \frac{e^{-4}4^2}{2!} = 8e^{-4}
\end{aligned}
$$

So the probability of two or fewer visitors in a minute is

$$e^{-4} + 4e^{-4} + 8e^{-4} = 0.238.$$

Example

(ii) If the average number of visitors in 1 minute is 4, the average in 30 seconds is 2.

Poisson Distribution

Example

(ii) If the average number of visitors in 1 minute is 4, the average in 30 seconds is 2.
So for this example, our parameter $\lambda = 2$. So

$$P(X = 2) = \frac{e^{-2}2^2}{2!} = 2e^{-2} = 0.271.$$

The previous example is a standard example of a queueing process. These are very important in many applications in contemporary communications engineering. Other examples of this type include the number of calls to an exchange, the arrival of customers at a service desk or the arrival of cars at a toll booth.

Poisson Distribution - Mean and Variance

The mean and variance of a Poisson random variable with parameter λ are both equal to λ.

$$E(X) = \lambda, \qquad V(X) = \lambda.$$

Example

It is believed that the number of bookings taken per hour at an online travel agency follows a Poisson distribution. Past records indicate that the hourly number of bookings has a mean of 15 and a standard deviation of 2.5. Comment on the suitability of the Poisson distribution for this example?

Poisson Distribution - Mean and Variance

Example

If the number of hourly bookings at this travel agent did follow a Poisson distribution,. we would expect that the mean and variance should be equal. However, in this case

$$E(X) = 15, \quad V(X) = (2.5)^2 = 6.25.$$

This suggests that the Poisson distribution is not appropriate for this case.

Poisson Approximation to Binomial Distribution

Suppose

1. $n \longrightarrow \infty$
2. $p \longrightarrow 0$ with np staying constant

Then, writing $\lambda := np$, it can be shown that the binomial probabilities $bin(k; n, p)$ tend to the Poisson probability $\frac{e^{-\lambda}\lambda^k}{k!}$.

Observation

Poisson probabilities can be used to approximate binomial probabilities when n is large and p is small

As a rule of thumb, this approximation is acceptable if $n \geq 20$ and $p \leq 0.05$. If $n \geq 100$ and $np \leq 10$, it is usually an excellent approximation.

Example

It is known that 3% of the circuit boards from a production line are defective. If a random sample of 120 circuit boards is taken from this production line, use the Poisson approximation to estimate the probability that the sample contains:

(i) Exactly 2 defective boards.

(ii) At least 2 defective boards.

Example

In this case, $n \geq 100$ and $np \leq 10$. Also, $\lambda = np = 120(0.03) = 3.6$.

(i) $P(X = 2) = \frac{e^{-3.6}(3.6)^2}{2!} = 0.177$.
Binomial calculation also gives an answer of 0.177 (0.1766)

Poisson Approximation to the Binomial Distribution

Example

In this case, $n \geq 100$ and $np \leq 10$. Also, $\lambda = np = 120(0.03) = 3.6$.

(i) $P(X = 2) = \frac{e^{-3.6}(3.6)^2}{2!} = 0.177$.
Binomial calculation also gives an answer of 0.177 (0.1766)

(ii)

$$\begin{aligned} P(X \geq 2) &= 1 - (P(X = 0) + P(X = 1)) \\ &= 1 - (0.027 + 0.097) = 0.876. \end{aligned}$$

Binomial distribution gives an answer of $1 - (0.026 + 0.096) = 0.878$.