Poisson distribution

The **Poisson distribution** is a discrete probability distribution that is most commonly used for for modeling situations in which we are counting the number of occurrences of an event in a particular interval of time where the occurrences are independent from one another and, on average, they occur at a given rate λ. That is, the Poisson distribution models independent events that occur randomly, but over a long enough period of time, the rate of occurrences converges to λ. For example, the probability distribution over the number of earthquakes in a year can be modeled as a Poisson distribution because although earthquakes occur randomly, over a long period of time, the number of earthquakes over time approaches a constant rate.

The concept of "rate" modeled by the Poisson does not need to be temporal. That is, it does not need to be the number of events per unit of time. Rather, the rate can measure the number of events per unit of length or area. For example, the number of chocolate chips in a cookie can be modeled with the Poisson distribution. On average, we might expect λ chocolate chips per unit of area of the cookie. However, chocolate chips are randomly dropped into the batter and thus, occurrences of chocolate chips in the cookie are random and independent from one another.

We denote a random variable X that follows a Poisson as

$$X \sim \text{Poiss}(\lambda)$$

where λ is called the "rate" parameter.

Definition 1 *A discrete random variable X follows a **Poisson distribution** if its probability mass function is given by*

$$p(x) := \frac{e^{-\lambda}\lambda^x}{x!}$$

*where λ is the **Poisson rate** parameter.*

Deriving the Poisson distribution

The Poisson distribution can be approximated by a binomial distribution for which the number of trials n is very large, and the probability of success p in a given trial is very small. For example, we can model the distribution over the number of earthquakes in a year as a binomial distribution in which each millisecond is a trial and the probability

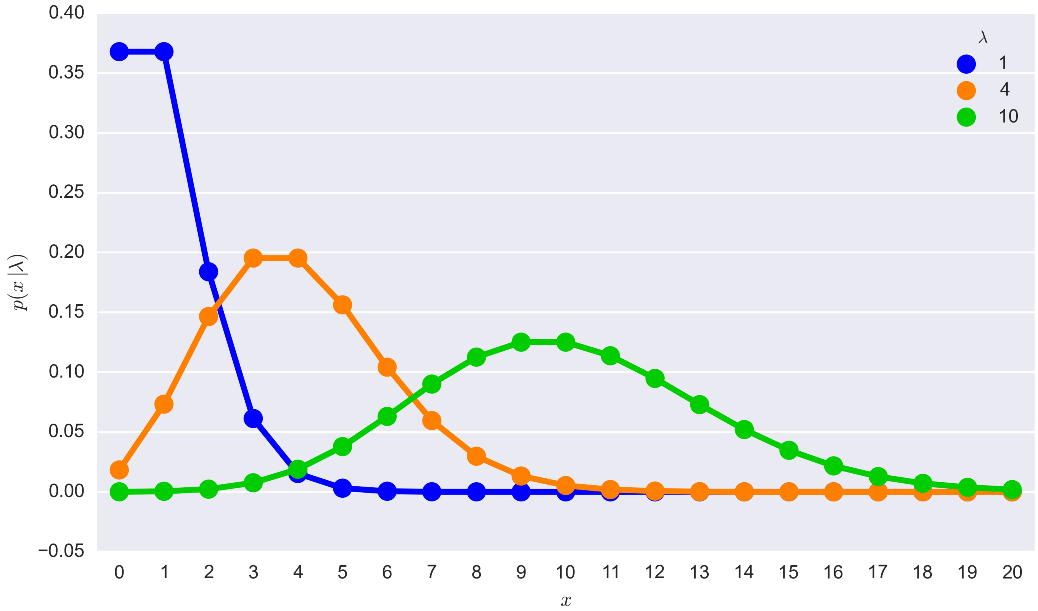

Figure 1: The probability mass function $p(x \mid \lambda)$ of the Poisson distribution for $\lambda = 1, 4, 10$.

of an earthquake in each millisecond is p. The number of trials n would then be the number of milliseconds in a year. Of course, the probability of an earthquake in a given millisecond is very small and the number of milliseconds in a given year is very large. On average, as a result, we get approximately λ earthquakes per year.

We will derive the Poisson by assuming some average rate λ and letting $\lambda := np$. When taking the limit, $n \to \infty$, we get the Poisson distribution.

Theorem 1 *Let $X \sim Bin(n, p)$ where $n \to \infty$, and $np = \lambda$. Then X follows a Poisson distribution.*

Proof:

The binomial p.m.f. is

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

Then,

$$\begin{aligned}
\binom{n}{x}p^x(1-p)^{n-x} &= \binom{n}{x}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x} && \text{Note 1}\\
&= \frac{n!}{(n-x)!x!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x} && \text{Note 2}\\
&= \frac{n!}{(n-x)!x!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n}\left(1-\frac{\lambda}{n}\right)^{-x}\\
&= \frac{n(n-1)(n-2)\dots(n-x+1)(n-x)!}{(n-x)!x!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n}\left(1-\frac{\lambda}{n}\right)^{-x}\\
&= \frac{n(n-1)(n-2)\dots(n-x+1)}{x!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n}\left(1-\frac{\lambda}{n}\right)^{-x} && \text{Note 3}\\
&= \left(\frac{\lambda^x}{x!}\right)\left(\frac{n(n-1)(n-2)\dots(n-x+1)}{n^x}\right)\left(1-\frac{\lambda}{n}\right)^{n}\left(1-\frac{\lambda}{n}\right)^{-x}\\
&= \left(\frac{\lambda^x}{x!}\right)\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\dots\left(\frac{n-x+1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{n}\left(1-\frac{\lambda}{n}\right)^{-x} && \text{Note 4}
\end{aligned}$$

Now we take the limit as n approaches infinity:

$$\begin{aligned}
&\lim_{n\to\infty}\left[\left(\frac{\lambda^x}{x!}\right)\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\dots\left(\frac{n-x+1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{n}\left(1-\frac{\lambda}{n}\right)^{-x}\right]\\
=&\left(\frac{\lambda^x}{x!}\right)\lim_{n\to\infty}\left(\frac{n}{n}\right)\lim_{n\to\infty}\left(\frac{n-1}{n}\right)\dots\lim_{n\to\infty}\left(\frac{n-x+1}{n}\right)\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n}\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-x} \quad \text{Note 5}\\
=&\frac{\lambda^x e^{-\lambda}}{x!}
\end{aligned}$$

Notes:

1. Substitute $p = \frac{\lambda}{n}$
2. Expand the binomial coefficient
3. Cancel the $(n-x)!$ from the numerator and denominator.
4. Note that in the numerator of the first term (i.e. $n(n-1)(n-2)\dots(n-x+1)$), there are x factors. This matches the number of n's being multiplied in the denominator.
5. By the algebraic limit theorem
6. Notice that $\lim_{n\to\infty}\left(\frac{n-i}{n}\right)$ approaches 1 for all i. Then, $\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}$. Lastly, $\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-x} = 1$.

$\square$

Mean

Theorem 2 *Given a random variable $X \sim Pois(\lambda)$, its mean is given by,*

$$E(X) = \lambda$$

Proof:

$$\begin{aligned}
E(X) &= \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} \\
&= \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} && \text{because the first term is zero} \\
&= e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!} \\
&= \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} && \text{pulling out a } \lambda \\
&= \lambda e^{-\lambda} e^{\lambda} && \text{we recognize the Taylor Series of } e^{\lambda} \\
&= \lambda
\end{aligned}$$

□

Variance

Theorem 3 *Given a random variable $X \sim Pois(\lambda)$, its variance is given by,*

$$Var(X) = \lambda$$

Proof:

$$
\begin{aligned}
\text{Var}(X) &= E\left(X^2\right) - E(X)^2 \\
&= E\left(X^2\right) - \lambda^2 && \text{Note 1} \\
&= \sum_{x=0}^{\infty} \left[x^2 \frac{e^{-\lambda}\lambda^x}{x!} \right] - \lambda^2 && \text{Note 2} \\
&= \sum_{x=1}^{\infty} \left[x^2 \frac{e^{-\lambda}\lambda^x}{x!} \right] - \lambda^2 && \text{Note 3} \\
&= e^{-\lambda} \sum_{x=1}^{\infty} \left[x^2 \frac{\lambda^x}{x!} \right] - \lambda^2 \\
&= e^{-\lambda} \sum_{x=1}^{\infty} \left[x \frac{\lambda^x}{(x-1)!} \right] - \lambda^2 \\
&= e^{-\lambda} \sum_{x=1}^{\infty} \left[(x-1)\frac{\lambda^x}{(x-1)!} + \frac{\lambda^x}{(x-1)!} \right] - \lambda^2 \\
&= e^{-\lambda} \left[\lambda^2 \sum_{x=1}^{\infty} \left((x-1)\frac{\lambda^{x-2}}{(x-1)!} \right) + \lambda \sum_{x=1}^{\infty} \left(\frac{\lambda^{x-1}}{(x-1)!} \right) \right] - \lambda^2 \\
&= e^{-\lambda} \left[\lambda^2 \sum_{x=2}^{\infty} \left((x-1)\frac{\lambda^{x-2}}{(x-1)!} \right) + \lambda \sum_{x=1}^{\infty} \left(\frac{\lambda^{x-1}}{(x-1)!} \right) \right] - \lambda^2 && \text{Note 4} \\
&= e^{-\lambda} \left[\lambda^2 \sum_{x=2}^{\infty} \left(\frac{\lambda^{x-2}}{(x-2)!} \right) + \lambda \sum_{x=1}^{\infty} \left(\frac{\lambda^{x-1}}{(x-1)!} \right) \right] - \lambda^2 \\
&= e^{-\lambda} \left[\lambda^2 \sum_{i=0}^{\infty} \left(\frac{\lambda^{i}}{i!} \right) + \lambda \sum_{j=0}^{\infty} \left(\frac{\lambda^{j}}{j!} \right) \right] - \lambda^2 && \text{Note 5} \\
&= e^{-\lambda} \left(\lambda^2 e^{\lambda} + \lambda e^{\lambda} \right) - \lambda^2 && \text{Note 6} \\
&= \lambda
\end{aligned}
$$

Notes:

1. by Theorem 2
2. by LOTUS
3. first term is zero
4. starting point of first summation can start from $x = 2$ because the first term is zero

5. let $i := x - 2$ and $j := x - 1$

6. Taylor series expansion of e^{λ}

$\square$