Poisson distribution

It is often useful to define a random variable that counts the number of events that occur within certain specified boundaries. For example, the average number of telephone calls received by customer service within a certain time limit. The Poisson distribution is often appropriate to model such situations.

Definition Poisson RV

A random variable X with a Poisson distribution takes the values $x = 0, 1, 2, \ldots$ with a probability mass function

$$pois(x;\mu) := P(X = x) = \frac{e^{-\mu}\,\mu^{x}}{x!}$$

where μ is the parameter of the distribution.[a]

[a] Some textbooks use λ for the parameter. We will use λ for the intensity of the Poisson process, to be discussed later.

Theorem: Mean and variance of Poisson RV

For Poisson RV with parameter μ,

$$\mathbb{E}(X) = V(X) = \mu.$$

Proof. Recall the Taylor series expansion of e^x:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$$

Now,

$$\mathbb{E}(X) = \sum x * pois(x,\mu) = \sum_{x=0}^{\infty} x\,\frac{e^{-\mu}\,\mu^{x}}{x!} = \sum_{x=1}^{\infty} \frac{x\,e^{-\mu}\,\mu\,\mu^{x-1}}{x(x-1)!} =$$

$$= \mu\,e^{-\mu}\sum_{x=1}^{\infty}\frac{\mu^{x-1}}{(x-1)!} = \mu\,e^{-\mu}\left[1 + \frac{\mu}{1!} + \frac{\mu^2}{2!} + \frac{\mu^3}{3!}\ldots\right] = \mu\,e^{-\mu}\,e^{\mu} = \mu$$

To find $\mathbb{E}(X^2)$, let us consider the factorial expression $\mathbb{E}[X(X-1)]$.

$$\mathbb{E}[X(X-1)] = \sum_{x=0}^{\infty} x(x-1)\,\frac{e^{-\mu}\,\mu^{x}}{x!} = \sum_{x=2}^{\infty} x(x-1)\,\frac{\mu^2\,e^{-\mu}\,\mu^{x-2}}{x(x-1)(x-2)!}$$

$$= \mu^2\,e^{-\mu}\sum_{x=2}^{\infty}\frac{\mu^{x-2}}{(x-2)!} = \mu^2\,e^{-\mu}\,e^{\mu} = \mu^2$$

Therefore, $\mathbb{E}[X(X-1)] = \mathbb{E}(X^2) - \mathbb{E}(X) = \mu^2$. Now we can solve for $\mathbb{E}(X^2)$ which is $\mathbb{E}(X^2) = \mathbb{E}[X(X-1)] + \mathbb{E}(X) = \mu^2 + \mu$.

Thus,

$$V(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \mu^2 + \mu - \mu^2 = \mu.$$

□

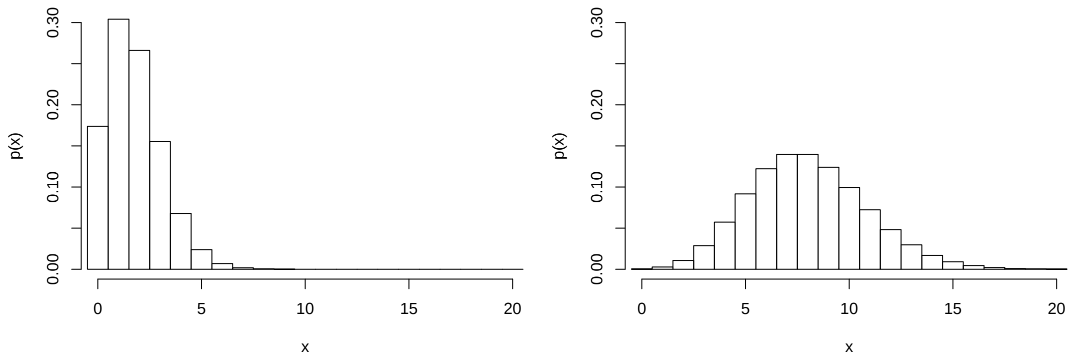

Figure 3.5: Poisson PMF: left, with $\mu = 1.75$; right, with $\mu = 8$

Example 1.
During World War II, the Nazis bombed London using V-2 missiles. To study the locations where missiles fell, the British divided the central area of London into 576 half-kilometer squares.[i] The following is the distribution of counts per square

Number of missiles in a square	Number of squares	Expected (Poisson) Number of squares
0	229	227.5
1	211	211.3
2	93	98.1
3	35	30.4
4	7	7.1
5 and over	1	1.6
Total	576	576.0

Are the counts suggestive of Poisson distribution?

Solution. The total number of missiles is $1(211) + 2(93) + 3(35) + 4(7) + 5(1) = 535$ and the average number per square, $\mu = 0.9288$. If the Poisson distribution holds, then the expected number of 0 squares (out of 576) will be

$$576 \times P(X = 0) = 576 \times \frac{e^{-0.9288}\, 0.9288^0}{0!} = 227.5$$

The same way, fill out the rest of the expected counts column. As you can see, the data match the Poisson model very closely!

Poisson distribution is often mentioned as a distribution of *spatial randomness.* As a result, British command were able to conclude that the missiles were unguided. □

Using the CDF

Knowledge of CDF (cumulative distribution function) is useful for calculating probabilities of the type $P(a \le X \le b)$. In fact,

$$P(a < X \le b) = F_X(b) - F_X(a) \tag{3.2}$$

(you have to carefully watch strict and non-strict inequalities). We might use CDF tables (see Appendix) to calculate such probabilities. Nowadays, CDF's of popular distributions are built into various software packages.

Example 2.
During a laboratory experiment, the average number of radioactive particles passing through a counter in one millisecond is 4. What is the probability that 6 particles enter the counter in a given millisecond? What is the probability of **at least** 6 particles?

Solution. Using the Poisson distribution with $x = 6$ and $\mu = 4$, we get

$$pois(6;4) = \frac{e^{-4}4^6}{6!} = 0.1042$$

Alternatively, using the CDF, $P(X = 6) = P(5 < X \leq 6) = F(6) - F(5)$. Using the Poisson table, $P(X = 6) = 0.8893 - 0.7851 = 0.1042$.

To find $P(X \geq 6)$, use $P(5 < X \leq \infty) = F(\infty) - F(5) = 1 - 0.7851 = 0.2149$ □

Poisson approximation for Binomial

Poisson distribution was originally derived as a limit of Binomial when $n \to \infty$ while $p = \mu/n$, with fixed μ. We can use this fact to estimate Binomial probabilities for large n and small p.

Example 3
At a certain industrial facility, accidents occur infrequently. It is known that the probability of an accident on any given day is 0.005 and the accidents are independent of each other. For a given period of 400 days, what is the probability that
(a) there will be an accident on only one day?
(b) there are at most two days with an accident?

Solution. Let X be a binomial random variable with $n = 400$ and $p = 0.005$. Thus $\mu = np = (400)(0.005) = 2$. Using the Poisson approximation,

a) $P(X = 1) = \frac{e^{-2}\,2^1}{1!} = 0.271$

b) $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = \frac{e^{-2}\,2^0}{0!} + \frac{e^{-2}\,2^1}{1!} + \frac{e^{-2}\,2^2}{2!}$
$= 0.1353 + 0.271 + 0.271 = 0.6766$ □

Exercises

1.
Number of cable breakages in a year is known to have Poisson distribution with $\mu = 0.32$.

a) Find the mean and standard deviation of the number of cable breakages in a year.

b) According to Chebyshev's inequality, what is the upper bound for $P(X \geq 2)$?

c) What is the exact probability $P(X \geq 2)$, based on Poisson model?

2.
Bolted assemblies on a hull of spacecraft may become loose with probability 0.005. There are 96 such assemblies on board. Assuming that assemblies behave statistically independently, find the probability that there is at most one loose assembly on board.