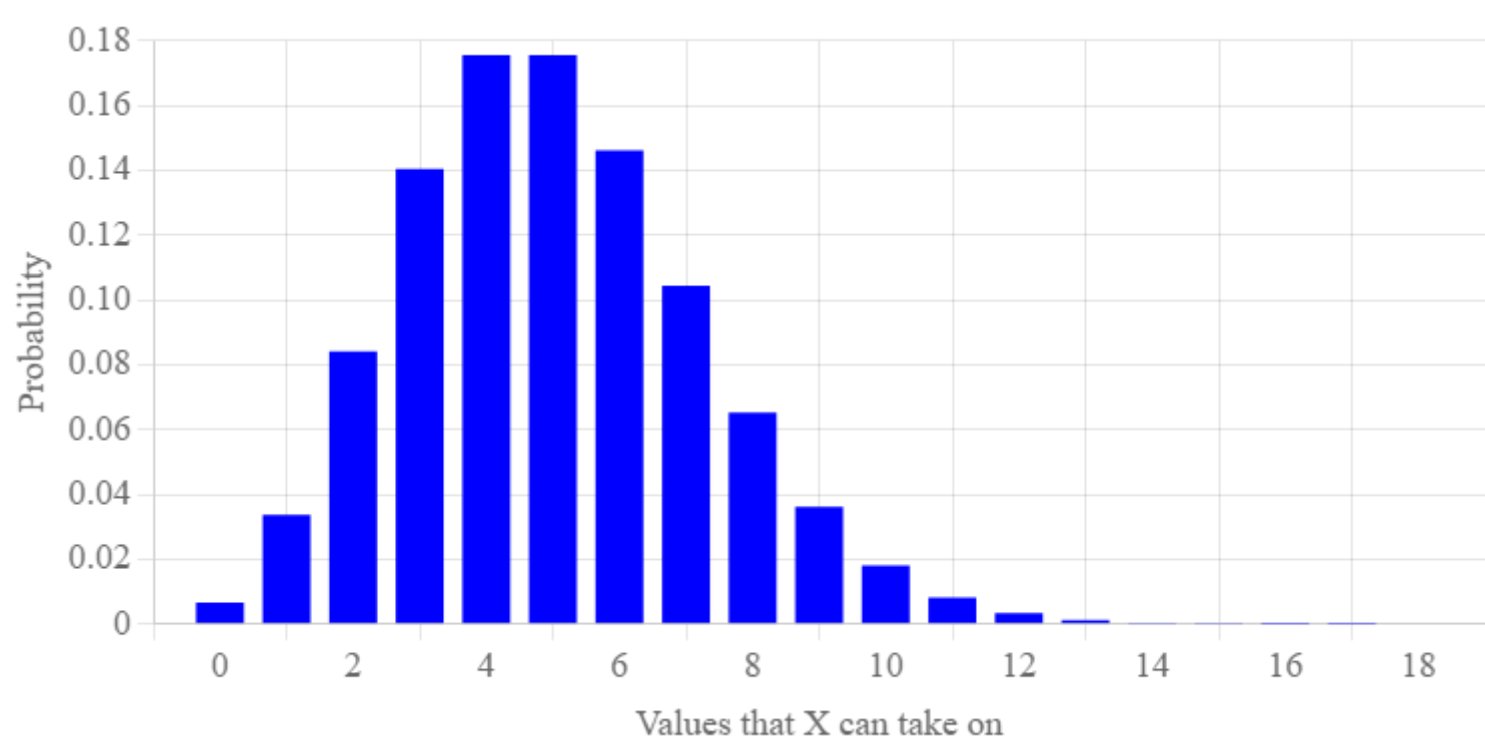# Poisson Distribution

A Poisson random variable gives the probability of a given number of events in a fixed interval of time (or space). It makes the Poisson assumption that events occur with a known constant mean rate and independently of the time since the last event.

Poisson Random Variable

Notation:	$X \sim \text{Poi}(\lambda)$
Description:	Number of events in a fixed time frame if (a) the events occur with a constant mean rate and (b) they occur independently of time since last event.
Parameters:	$\lambda \in \{0, 1, \ldots\}$, the constant average rate.
Support:	$x \in \{0, 1, \ldots\}$
PMF equation:	$P(X = x) = \dfrac{\lambda^x e^{-\lambda}}{x!}$
Expectation:	$\text{E}[X] = \lambda$
Variance:	$\text{Var}(X) = \lambda$

**PMF graph:**

Parameter λ: `5`

Poisson Intuition

In this section we show the intuition behind the Poisson derivation. It is both a great way to deeply understand the Poisson, as well as good practice with Binomial distributions.

Let's work on the problem of predicting the chance of a given number of events occurring in a fixed time interval — the next minute. For example, imagine you are working on a ride sharing application and you care about the probability of how many requests you get from a particular area. From historical data, you know that the average requests per minute is $\lambda = 5$. What is the probability of getting 1, 2, 3, etc requests in a minute?

💡: We could approximate a solution to this problem by using a binomial distribution! Lets say we split our minute into 60 seconds, and make each second an [indicator Bernoulli](#) variable — you either get a request or you don't. If you get a request in a second, the indicator is 1. Otherwise it is 0. Here is a visualization of our 60 binary-indicators. In this example imagine we have requests at 2.75 and 7.12 seconds. the corresponding indicator variables are **blue** filled in boxes:
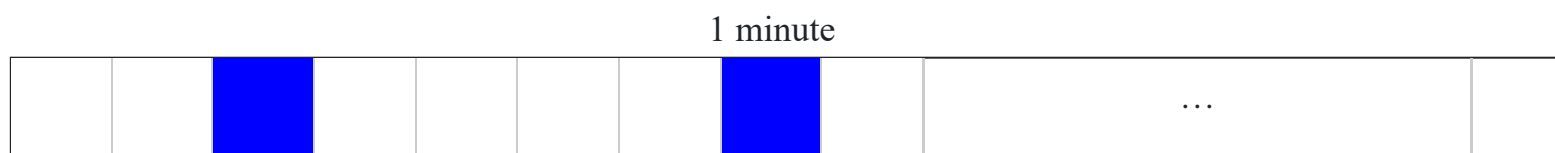
1 minute

The total number of requests received over the minute can be approximated as the sum of the sixty indicator variables, which conveniently matches the description of a [binomial](#) — a sum of Bernoullis. Specifically define X to be the number of requests in a minute. X is a binomial with $n = 60$ trials. What is the probability, p, of a success on a single trial? To make the expectation of X equal the observed historical average $\lambda = 5$ we should choose p so that $\lambda = \mathrm{E}[X]$.

$$
\begin{aligned}
\lambda &= \mathrm{E}[X] && \text{Expectation matches historical average} \\
\lambda &= n \cdot p && \text{Expectation of a Binomial is } n \cdot p \\
p &= \frac{\lambda}{n} && \text{Solving for } p
\end{aligned}
$$

In this case since $\lambda = 5$ and $n = 60$, we should choose $p = 5/60$ and state that $X \sim \mathrm{Bin}(n = 60, p = 5/60)$. Now that we have a form for X we can answer probability questions about the number of requests by using the Binomial PMF:

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

So for example:

$$
\mathrm{P}(X = 1) = \binom{60}{1}(5/60)^1(55/60)^{60-1} \approx 0.0295
$$

$$
\mathrm{P}(X = 2) = \binom{60}{2}(5/60)^2(55/60)^{60-2} \approx 0.0790
$$

$$
\mathrm{P}(X = 3) = \binom{60}{3}(5/60)^3(55/60)^{60-3} \approx 0.1389
$$

Great! But don't forget that this was an approximation. We didn't account for the fact that there can be more than one event in a single second. One way to assuage this issue is to divide our minute into more fine-grained intervals (the choice to split it into 60 seconds was rather arbitrary). Instead lets divide our minute into 600 deciseconds, again with requests at 2.75 and 7.12 seconds:

1 minute
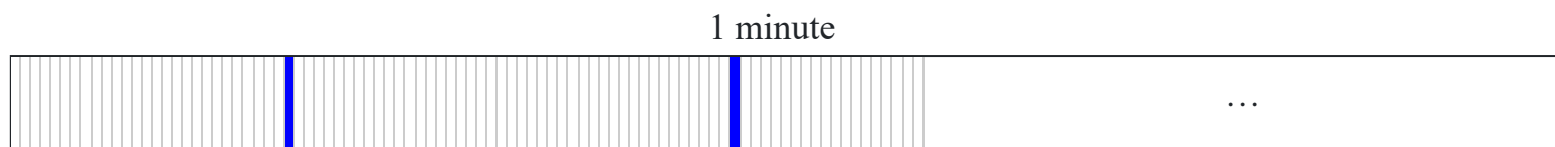
Now $n = 600$, $p = 5/600$ and $X \sim \mathrm{Bin}(n = 600, p = 6/600)$. We can repeat our example calculations using this better approximation:

$$
\mathrm{P}(X = 1) = \binom{600}{1}(5/600)^1(595/60)^{600-1} \approx 0.0333
$$

$$
\mathrm{P}(X = 2) = \binom{600}{2}(5/600)^2(595/600)^{600-2} \approx 0.0837
$$

$$
\mathrm{P}(X = 3) = \binom{600}{3}(5/600)^3(595/600)^{600-3} \approx 0.1402
$$

Choose any value of n, the number of buckets to divide our minute into: 0

The larger n is, the more accurate the approximation. So what happens when n is infinity? It becomes a Poisson!

Poisson, a Binomial in the limit

Or if we really cared about making sure that we don't get two events in the same bucket, we can divide our minute into infinitely small buckets:
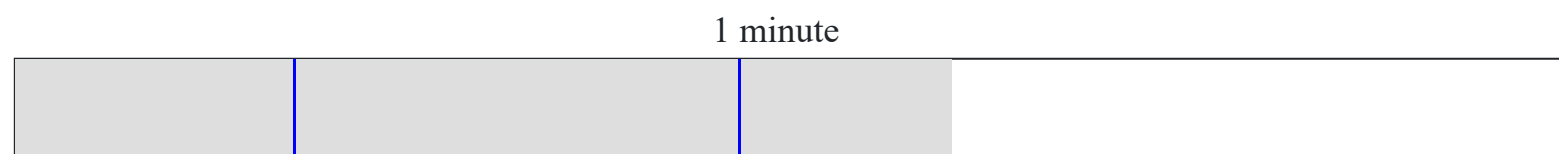
1 minute

Proof: Derivation of the Poisson

What does the PMF of X look like now that we have infinite divisions of our minute? We can write the equation and think about it as n goes to infinity. Recall that p still equals λ/n:

$$P(X = x) = \lim_{n \to \infty} \binom{n}{x} (\lambda/n)^x (1 - \lambda/n)^{n-x}$$

While it may look intimidating, this expression simplifies nicely. This proof uses a few special limit rules that we haven't introduced in this book:

$$
\begin{aligned}
P(X = x) &= \lim_{n \to \infty} \binom{n}{x} (\lambda/n)^x (1 - \lambda/n)^{n-x} && \text{Start: binomial in the limit} \\[2mm]
&= \lim_{n \to \infty} \binom{n}{x} \cdot \frac{\lambda^x}{n^x} \cdot \frac{(1 - \lambda/n)^n}{(1 - \lambda/n)^x} && \text{Expanding the power terms} \\[2mm]
&= \lim_{n \to \infty} \frac{n!}{(n-x)!x!} \cdot \frac{\lambda^x}{n^x} \cdot \frac{(1 - \lambda/n)^n}{(1 - \lambda/n)^x} && \text{Expanding the binomial term} \\[2mm]
&= \lim_{n \to \infty} \frac{n!}{(n-x)!x!} \cdot \frac{\lambda^x}{n^x} \cdot \frac{e^{-\lambda}}{(1 - \lambda/n)^x} && \text{Rule } \lim_{n \to \infty} (1 - \lambda/n)^n = e^{-\lambda} \\[2mm]
&= \lim_{n \to \infty} \frac{n!}{(n-x)!x!} \cdot \frac{\lambda^x}{n^x} \cdot \frac{e^{-\lambda}}{1} && \text{Rule } \lim_{n \to \infty} \lambda/n = 0 \\[2mm]
&= \lim_{n \to \infty} \frac{n!}{(n-x)!} \cdot \frac{1}{x!} \cdot \frac{\lambda^x}{n^x} \cdot \frac{e^{-\lambda}}{1} && \text{Splitting first term} \\[2mm]
&= \lim_{n \to \infty} \frac{n^x}{1} \cdot \frac{1}{x!} \cdot \frac{\lambda^x}{n^x} \cdot \frac{e^{-\lambda}}{1} && \lim_{n \to \infty} \frac{n!}{(n-x)!} = n^x \\[2mm]
&= \lim_{n \to \infty} \frac{\lambda^x}{x!} \cdot \frac{e^{-\lambda}}{1} && \text{Cancel } n^x \\[2mm]
&= \frac{\lambda^x \cdot e^{-\lambda}}{x!} && \text{Simplify}
\end{aligned}
$$

That is a beautiful expression! Now we can calculate the real probability of number of requests in a minute, if the historical average is $\lambda = 5$:

$$P(X = 1) = \frac{5^1 \cdot e^{-5}}{1!} = 0.03369$$

$$P(X = 2) = \frac{5^2 \cdot e^{-5}}{2!} = 0.08422$$

$$P(X = 3) = \frac{5^3 \cdot e^{-5}}{3!} = 0.14037$$

This is both more accurate and much easier to compute!

Changing time frames

Say you are given a rate over one unit of time, but you want to know the rate in another unit of time. For example, you may be given the rate of hits to a website per minute, but you want to know the probability over a 20 minute period. You would just need to multiply this rate by 20 in order to go from the "per 1 minute of time" rate to obtain the "per 20 minutes of time" rate.