Lecture 16: The Poisson Distribution

Anup Rao

February 12, 2018

We introduce and motivate the Poisson Distribution.

WE HAVE BEEN PLAYING with a number of different distributions in this class. Here are some of them:

Bernoulli Toss a coin that is heads with probability p. Let X be the outcome of the coin toss. Then X has a Bernoulli distribution.

Binomial Toss n coins independently, each of which gives heads with probability p. Let X be the number of heads. Then X has a binomial distribution.

Geometric Toss an infinite number of coins in sequence, each of which is heads with probability p. Let X be the number of coin tosses before you see the first heads. X has a geometric distribution.

As you can see, we have been obsessed with coins in this course. That's mainly because the Bernoulli distribution and its derivatives are extremely nice and clean to work with. They do seem to give a good approximation in many real-world applications as well. However, there are many other kinds of distributions that are better suited to understanding the world.

The *Poisson process* is one such extremely natural distribution:

$$p_\lambda(k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!},$$

for $k = 0, 1, 2, \ldots$.

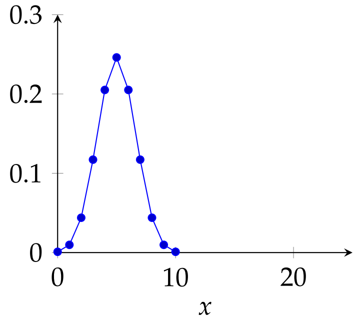
It is easy to see that this is actually a distribution on the non-negative integers. Recall the Taylor series expansion:

$$e^\lambda = 1 + \lambda + \lambda^2/2! + \lambda^3/3! + \ldots.$$

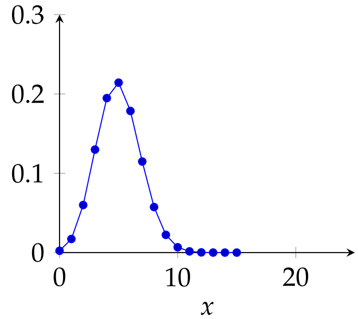
We can use this expansion to prove:

$$\sum_{k=0}^{\infty} p_\lambda(k) = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} \cdot e^\lambda = 1.$$

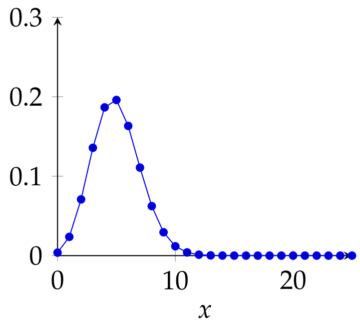
The distribution p_λ is actually just as natural as the other distributions we have discussed above. Indeed, it corresponds to the *limit of the Binomial distribution as $n \to \infty$*.



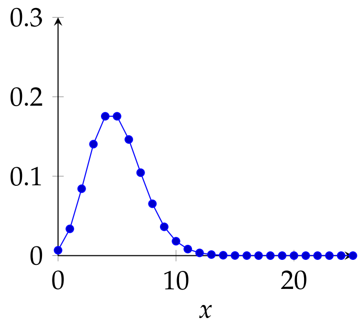
(a) Binomial with $n = 10$, $p = 1/2$.



(b) Binomial with $n = 15$, $p = 1/3$.



(c) Binomal with $n = 25$, $p = 1/5$.



(d) Poisson with $\lambda = 5$.

Figure 1: Three binomial distributions are shown, each of which has expected value 5. The last figure shows the Poisson distribution with $\lambda = 5$. Note how as n gets larger and larger, the Binomial converge to the Poisson distribution.

To see this, suppose X is the binomial distribution induced by n coin tosses, each of which is heads with probability p. Let $\lambda = pn$ be the expected number of heads. Then we have

$$\begin{aligned} p(X = k) &= \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \\ &= \frac{n \cdot (n-1) \dots (n-k+1)}{k!} \cdot \left(\frac{\lambda}{n}\right)^k \cdot \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{n}{n} \cdot \frac{n-1}{n} \dots \frac{n-k+1}{n} \cdot \frac{\lambda^k}{k!} \cdot \left(1 - \frac{\lambda}{n}\right)^{n-k}. \end{aligned}$$

since $p = \lambda/n$

Now, if we hold λ and k constant, and let $n \to \infty$, we have

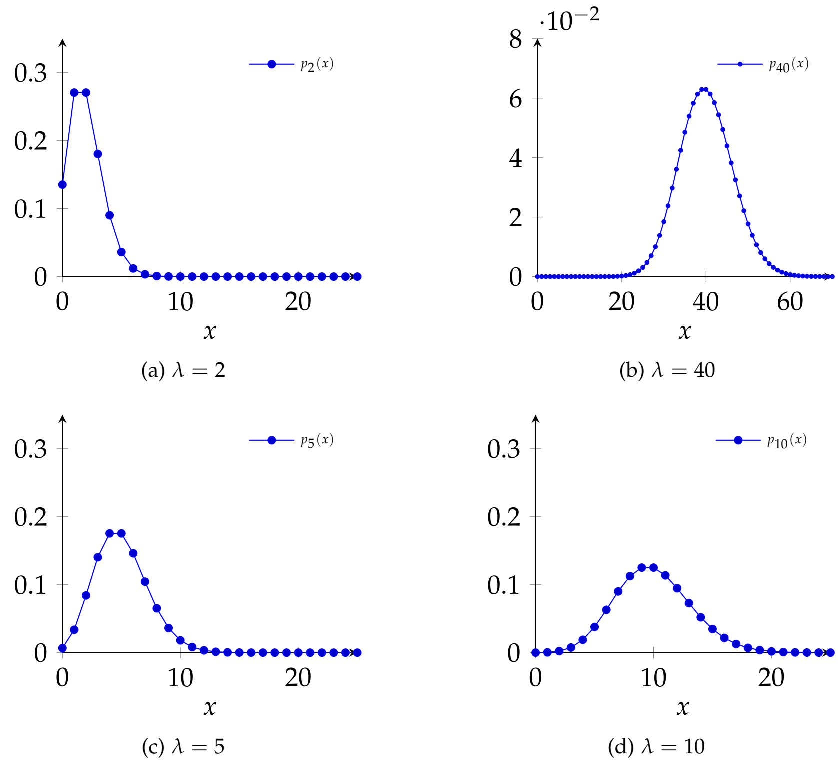
$$\frac{n}{n} \cdot \frac{n-1}{n} \dots \frac{n-k+1}{n} \to 1 \qquad \left(1 - \frac{\lambda}{n}\right)^{-k} \to 1 \qquad \left(1 - \frac{\lambda}{n}\right)^{n} \to e^{-\lambda},$$

Recall that $e = \lim_{\frac{n}{\lambda} \to \infty} (1 - \lambda/n)^{n/\lambda}$

so we get

$$p(X = k) \to \frac{\lambda^k}{k!} \cdot e^{-\lambda} = p_\lambda(k).$$

So, the Poisson distribution is basically the distribution of the number of heads when you toss an *infinitely many coins* and you expect to see λ heads.



(a) $\lambda = 2$ (b) $\lambda = 40$ (c) $\lambda = 5$ (d) $\lambda = 10$

Figure 2: The Poisson distribution, for various choices of λ. Note that the y axis has a different range in each figure. The large values of λ actually have much smaller probabilities.

By construction, if X is distributed according the Poisson distribution above, then

$$\mathbb{E}[X] = \lambda,$$

since the expected number of heads is always λ.

We can also calculate $\mathbb{E}[X]$ explicitly:

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{k=0}^{\infty} p_\lambda(k) \cdot k \\
&= \sum_{k=1}^{\infty} p_\lambda(k) \cdot k \qquad \text{since the contribution when } k = 0 \text{ is } 0 \\
&= \sum_{k=1}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k \cdot k}{k!} \\
&= \lambda e^{-\lambda} \cdot \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda \cdot e^{-\lambda} \cdot e^{\lambda} = \lambda.
\end{aligned}$$

What about $\mathrm{Var}[X]$? If Y was binomial, then we calculated that $\mathrm{Var}[Y] = n(p - p^2) = np - np^2 = \lambda - p\lambda$. As we take $n \to \infty$ and hold λ constant we have $p \to 0$, so we should get

$$\mathrm{Var}[X] = \lambda.$$

We can also calculate it directly:

$$
\begin{aligned}
\mathbb{E}\left[X^2\right] &= \sum_{k=0}^{\infty} p_\lambda(k) \cdot k^2 \\
&= \sum_{k=1}^{\infty} p_\lambda(k) \cdot k^2 && \text{since the contribution when } k = 0 \text{ is } 0 \\
&= \sum_{k=1}^{\infty} e^{-\lambda} \cdot \frac{\lambda^k \cdot k^2}{k!} \\
&= \lambda e^{-\lambda} \cdot \sum_{k=1}^{\infty} \frac{\lambda^{k-1} \cdot k}{(k-1)!} \\
&= \lambda e^{-\lambda} \cdot \left(\sum_{k=1}^{\infty} \frac{\lambda^{k-1} \cdot (k-1)}{(k-1)!} + \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \right) \\
&= \lambda e^{-\lambda} \cdot \left(\lambda \cdot \sum_{k=2}^{\infty} \frac{\lambda^{k-2} \cdot (k-2)}{(k-1)!} + e^{\lambda} \right) \\
&= \lambda e^{-\lambda} (\lambda e^{\lambda} + e^{\lambda}) = \lambda^2 + \lambda.
\end{aligned}
$$

the first term of the first sum is 0, and the second sum is just e^{λ}.

So:

$$\mathrm{Var}\left[X\right] = \mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Example: Modeling the Arrival of Traffic

To motivate why the Poisson distribution is so natural, consider the problem to trying to model the number of cars that pass through a given traffic intersection in τ hours. How should we model this?

It is very natural to enforce that the distribution should satisfy that the expected number of cars passing through the intersection should be proportional to τ. So, let us assume that this expectation is $\lambda \cdot \tau$, for some parameter λ. If $\lambda \cdot \tau \ll 1$, an approach might be to model the number of cars passing through as a Bernoulli with probability $\lambda \cdot \tau$. Then the expectation is correct, and maybe this makes sense. The problem is that this doesn't allow for the possibility that the number of cars is ever more than 1.

The next idea is to think of breaking up the interval of length τ into τ/δ pieces of length $\delta \ll \tau$, and think of the number of cars passing during each smaller interval as a Bernoulli. In order to make this work, we want the overall expected number of cars to still be $\lambda\tau$, so the probability of a car passing in each interval of length δ should be $\lambda\delta$. What happens when we take $\delta \to 0$? Then we see that the distribution on the number of cars in the interval τ behaves exactly like a Poisson process with probability $\lambda \cdot \tau$.