Lecture 19: Variance and Expectation of the Exponential Distribution, and the Normal Distribution

Anup Rao

May 15, 2019

Last time we defined the exponential random variable. This the time of the first arrival in the Poisson process with parameter λ. Recall that we computed its pdf to be

$$f(t) = \lambda e^{-\lambda t},$$

and its cdf to be

$$F(t) = 1 - e^{-\lambda t}.$$

Let us compute the variance and expectation of the exponential random variable. To compute the expectation, recall that the Poisson process is the limit of binomial distributions. We can think of it as n Bernoullis in each unit of time, each with parameter p, such that $pn = \lambda$. Now, we have computed the expected number of throws before we see the first success in the Bernoullis (this is the geometric random variable). The expectation is $1/p$. So, we should expect to see the first arrival after $1/p = n/\lambda$ throws. This corresponds to $1/\lambda$ units of time.

More formally, the expectation of the exponential random variable is then (using integration by parts)

$$\begin{aligned}
\mathbb{E}[X] &= \int_0^\infty t\lambda e^{-\lambda t}\, dt \\
&= -t \cdot e^{-\lambda t}\Big|_0^\infty - \int_0^\infty (-e^{-\lambda t})\, dt \\
&= 0 + (-1/\lambda)e^{-\lambda t}\Big|_0^\infty = 1/\lambda.
\end{aligned}$$

Next, let us compute the variance. As before, we could calculate this using the analogy to the binomial distribution. We computed the variance of the geometric random variable in Lecture 15, it was $\frac{1-p}{p^2}$, which is the same as $\frac{1-\lambda/n}{(\lambda/n)^2} = \frac{1-\lambda/n}{(\lambda/n)^2}$. This gives a standard deviation of

$$\sqrt{\frac{1-\lambda/n}{(\lambda/n)^2}} = \frac{n}{\lambda} \cdot \sqrt{1-\lambda/n}.$$

As n becomes larger, this converges to n/λ, or $1/\lambda$ units of time. So, the variance is $1/\lambda^2$. Similarly, the variance can be calculated by computing (using the product rule twice):

$$\begin{aligned}
\mathbb{E}\left[X^2\right] &= \int_0^\infty t^2 \lambda e^{-\lambda t}\, dt \\
&= -t^2 \cdot e^{-\lambda t}\Big|_0^\infty - \int_0^\infty (-2te^{-\lambda t})\, dt \\
&= 0 + (-2t/\lambda)e^{-\lambda t}\Big|_0^\infty - \int_0^\infty (-2/\lambda)e^{-\lambda t}\, dt \\
&= 0 + (-2/\lambda^2)e^{-\lambda t}\Big|_0^\infty \\
&= 2/\lambda^2.
\end{aligned}$$

So, the variance is $\mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right]^2 = 1/\lambda^2$, and the standard deviation is $1/\lambda$.

Fact 1 (Memorylessness)**.** *If X is distributed according to the exponential distribution, then $p(X > s + t | X > t) = p(X > s)$.*

Proof.

$$\begin{aligned}
p(X > s + t | X > t) &= \frac{p(X > s + t, X > t)}{p(X > t)} \\
&= \frac{p(X > s + t)}{p(X > t)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-\lambda s},
\end{aligned}$$

which is exactly the same as the probability that $X > s$. □

In homework, you will prove the following fact:

Fact 2. *If X and Y are independent Poisson's with parameters λ_1 and λ_2, then $X + Y$ is a Poisson with parameter $\lambda_1 + \lambda_2$.*

The most important distribution in the whole world

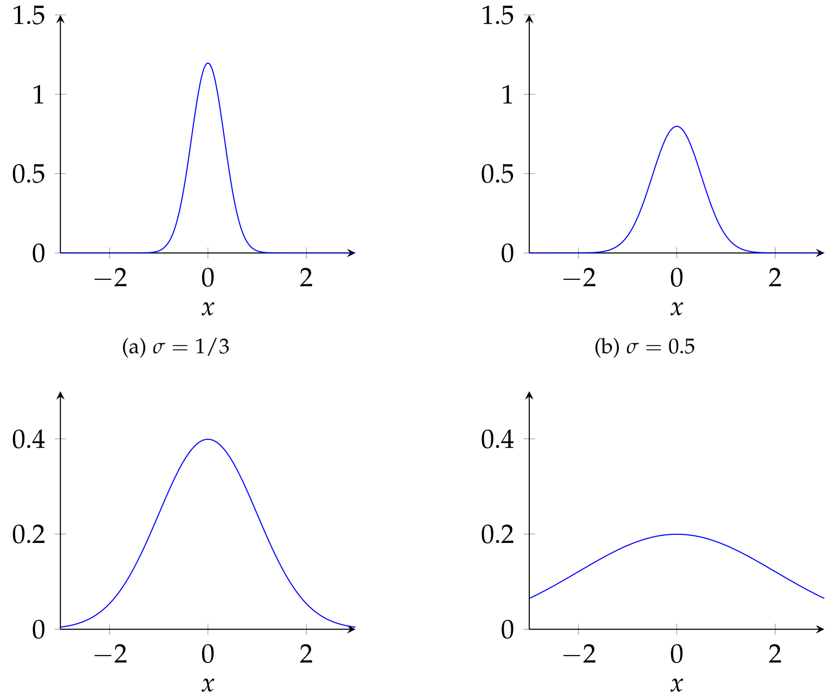
There is one distribution that is more important than all the others. It seems to be the right model for all kinds of processes observed in practice. It is called the *normal* distribution. The normal distribution is sometimes referred to as the *Gaussian* distribution.

The *central limit theorem*, which we discuss soon, provides a mathematical explanation for why the normal distribution is so commonly found in the wild.

The pdf of the normal distribution is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

where here μ and σ are parameters of the distribution. The formula has been set up so that μ is the expected value, and σ is the standard deviation of the normal.



(a) $\sigma = 1/3$ (b) $\sigma = 0.5$ (c) $\sigma = 1$ (d) $\sigma = 2$

Figure 1: The pdf of the normal distribution with $\mu = 0$.

It is a little bit tricky to check that the pdf of the normal distribution is a valid pdf, namely that

$$\int_{-\infty}^{\infty} f(x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} = 1.$$

See `https://en.wikipedia.org/wiki/Gaussian_integral`.

We do not do it here.

It is easy to see that μ is the expected value of the normal—the pdf is symmetric around μ. The value of the pdf at $\mu + \epsilon$ is equal to its value at $\mu - \epsilon$, so the average value must be μ. To compute the variance, we can first set $\mu = 0$, which doesn't change the variance. Then we have:

$$\mathbb{E}\left[X^2\right] = \int_{-\infty}^{\infty} x^2 f(x)\, dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x^2 \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx.$$

The integral can be evaluated using integration by parts:

$$\int_{-\infty}^{\infty} (x) \cdot (xe^{-\frac{x^2}{2\sigma^2}})\, dx = x \cdot (-\sigma^2)e^{-\frac{x^2}{2\sigma^2}})\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} (-\sigma^2)e^{-\frac{x^2}{2\sigma^2}}\, dx$$

$$= \sigma^2 \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}}\, dx.$$

The first term is 0, since $xe^{-x^2/2\sigma^2}$ goes to 0 as x gets large.

So, we conclude that

$$\begin{aligned}\mathbb{E}\left[X^2\right] &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(x)\cdot(xe^{-\frac{x^2}{2\sigma^2}})\,dx\\ &= \sigma^2\cdot\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}e^{-\frac{x^2}{2\sigma^2}}\,dx\\ &= \sigma^2.\end{aligned}$$

since the second term is the area under the pdf of the normal, which is 1.

Thus, the variance is (using $\mathbb{E}[X]=0$),

$$\begin{aligned}\mathrm{Var}[X] = \mathbb{E}\left[X^2\right] &= \sigma^2\cdot\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}e^{-\frac{x^2}{2\sigma^2}}\,dx\\ &= \sigma^2,\end{aligned}$$

since the integral of the pdf is 1.

If X is a normal with mean μ and standard deviation σ, then $aX+b$ is also normal, with mean $\mu+b$ and standard deviation $a\sigma$. Unfortunately, the cdf of the normal distribution has no nice closed form. However, it is not too hard to use programs to evaluate the cdf.

Theorem 3. *Given any real-valued distribution with expectation μ and standard deviation σ, suppose $X_1, X_2, \ldots, X_n$ are sampled independently according to this distribution and*

$$Y_n = \frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Then the cdf of Y_n converges to the cdf of the standard normal, in the sense that for every α,

$$\lim_{n\to\infty} p(Y_n \le \alpha) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\alpha} e^{-x^2/2}\,dx.$$

Note that $\int_{-\infty}^{\alpha} e^{-x^2/2}\,dx$ is just the cdf of the normal distribution with mean 0 and standard deviation 1. So, the theorem asserts that the cdf of the sum convergers to the cdf of the standard normal, after we shift and scale it appropriately.