Lecture 19: The Normal Distribution

Anup Rao

February 21, 2018

We discuss the normal distribution and the weak law of large numbers.

We have begun discussing continuous random variables. Last time, we talked about how to define the expectation and variance of a continuous random variables. Just like with discrete random variables we have:

Fact 1 (Linearity of Expectation)**.** $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$.

Fact 2 (Variance of Linear Transformations)**.** $\text{Var}[aX+b] = a^2 \cdot \text{Var}[X]$.

Last time we discussed the exponential distribution. Here are some facts that you can easily prove about the exponential:

Fact 3 (Memorylessness)**.** *If X is distributed according to the exponential distribution, then $p(X > s+t | X > t) = p(X > s)$.*

Proof.

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

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

In section, you have seen the following fact:

Fact 4. *If X and Y are independent Poisson's with parameters λ and λ', then $X+Y$ is a Poisson with parameter $\lambda + \lambda'$.*

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

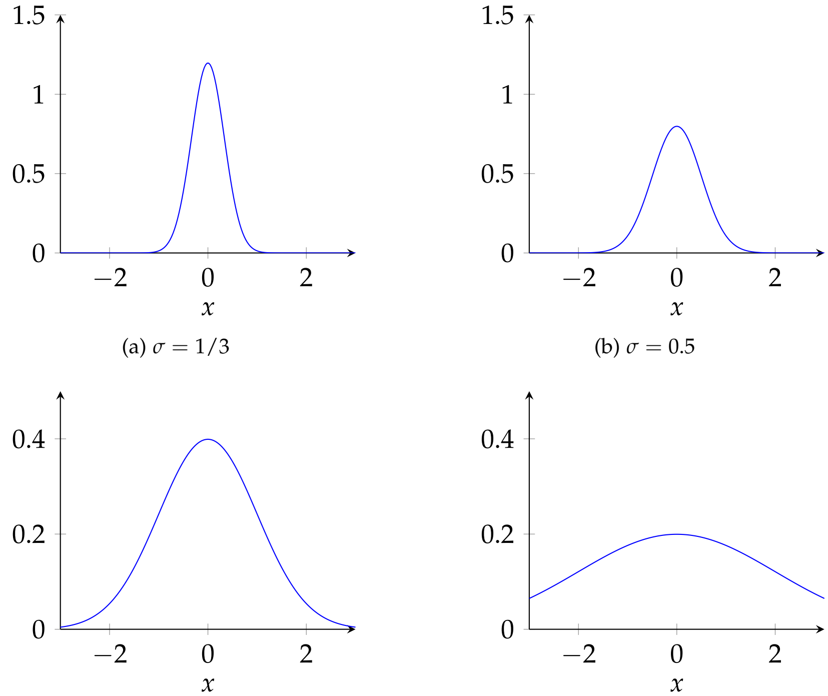
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

where here μ and σ are parameters of the distribution. The formula

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 here for a proof: `https://en.wikipedia.org/wiki/Gaussian_integral`



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

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

has been set up so that μ is the expected value, and σ is the standard deviation of the normal. 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:

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

The integral can be evaluated using integration by parts:

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

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

So, since $\mathbb{E}[X] = 0$,

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

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.

The Weak Law of Large Numbers

Often, we end up estimating the average of a random variable by taking the mean of a large number of samples. The weak law of large numbers says that this kind of experiment will produce a good estimate for the mean:

Theorem 5. *If $X_1, \ldots, X_n$ are independent and identically distributed, with expectation μ, and $Y_n = \frac{X_1+X_2+\ldots+X_n}{n}$, then for every $\epsilon > 0$,*

$$\lim_{n\to\infty} p(|Y_n - \mu| > \epsilon) = 0.$$

The weak law of large numbers is a straightforward consequence of Chebyshev's inequality. Basically, the variance of Y_n gets smaller and smaller as n gets larger, so the probability that Y_n deviates from the mean goes to 0.