In nature, many things - like height and weight of a human being - are distributed according to a normal distribution. There are many factors that contribute to its value, and the more complex something is the more it tends to resemble a bell curve. This is nuts - why should there be a single distribution that governs so many complicated things in the universe?

Central Limit Theorem, which together with the Law of Large Numbers are possibly the most profound statements in probability theory, gives us an insight as to why this is. It turns out that if you add together many identical variables, their sum starts looking more and more like a normal variable. we saw a glimpse of that when we looked at convolutions. Here are the results of adding several variables with uniform, exponential and normal densities:

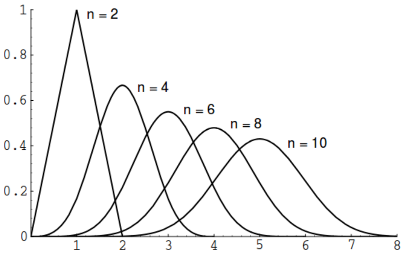


Figure 7.6: Convolution of n uniform densities.

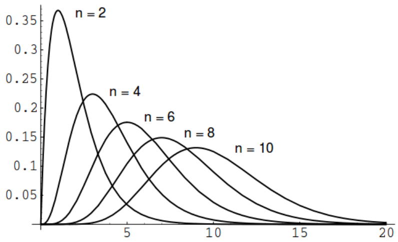


Figure 7.8: Convolution of n exponential densities with $\lambda = 1$.

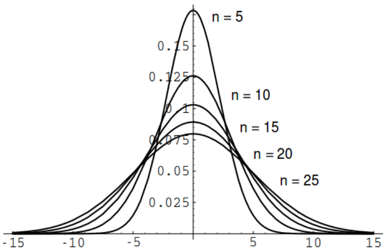


Figure 7.7: Convolution of n standard normal densities.

Adding together two normal variables doesn't make them any MORE like a normal variable - the sume is another normal variable. But the standard deviation, predictably, increases, so the sum is more spread out to the sides. Below is the precise statement of Central Limit Theorem.

Theorem 1 (Central Limit Theorem)

Let $S_n = X_1 + X_2 + \cdots + X_n$ be the either:

- *sum of n independent discrete random variables with common distribution $m(x)$, common mean μ and common variance σ^2.*
- *sum of n independent continuous random variables with common density $f(x)$, common mean μ and common variance σ^2.*

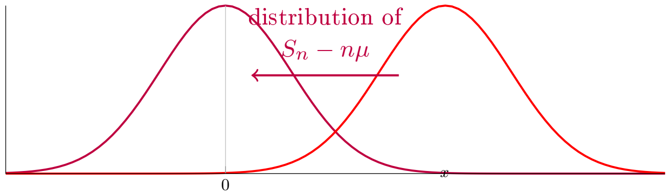
Let

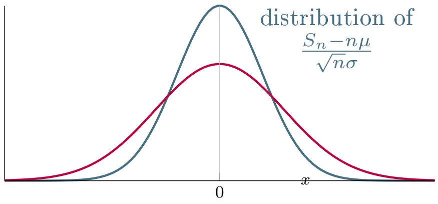
$$S_n^* = \frac{S_n - n\mu}{\sqrt{n}\sigma},$$

then for all $a < b$:

$$lim_{n\to\infty} P(a < S_n^* < b) = \frac{1}{2\pi}\int_a^b e^{-x^2/2}dx.$$

As you remember, $\frac{1}{2\pi}\int_a^b e^{-x^2/2}dx$ is the probability that a standard normal variable (i.e. with mean 0 and standard deviation 1) falls between a and b. S_n^* is called the *standardized* random variable. The only way we can say that two functions actually look similar, is with their values. So we need S_n to speak the same language as a standard normal variable, i.e. pretend that it's centered around 0 and is rescaled so that the standard deviation is 1. This is why we need S_n^*, it's the function that has the shape of S_n but has both of these properties.





The numbers a and b are values of the standardized variable that have their counterparts c and d in the original one:

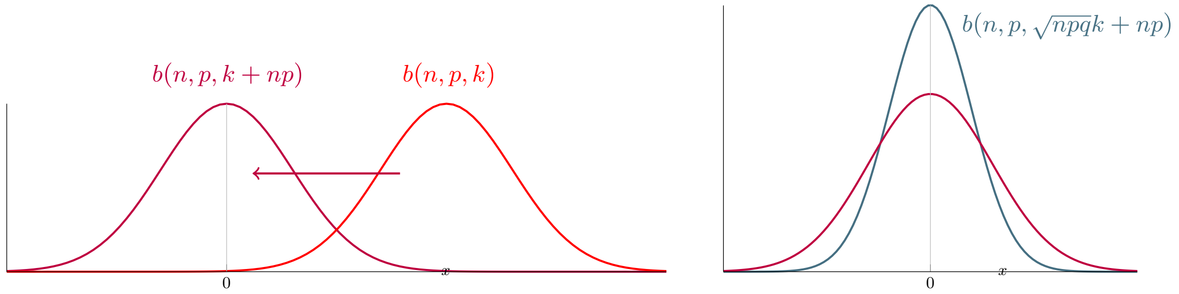
$$a = \frac{c - n\mu}{\sqrt{n}\sigma}, \quad b = \frac{d - n\mu}{\sqrt{n}\sigma}$$

The probability that S_n falls between c and d is the same as the probability that S_n^* falls between a and b. The CLT tells us that for large enough n, the probability that S_n falls between c and d can be estimated by the probability that a stadard normal variable falls between a and b, and that probability is

$$\frac{1}{2\pi}\int_a^b e^{-x^2/2}dx$$

and the numerical values can be found in the tables.

Example 1 (Bernoulli Trials)
Let S_n be the sum of n Bernoulli trials, each with probability of success p. Then each trial has mean $\mu = p$ and variance $\sigma = pq$. The distribution of S_n is described by the binomial distribution $b(n, p, k)$. Then The distribution of $S_n - np$ is described by $b(n, p, k + np)$. The distribution of $\frac{S_n - np}{\sqrt{npq}}$ is bescribed by $b(n, p, \sqrt{npq}k + np)$.



By CLT,

$$\lim_{n\to\infty} P(c < S_n < d) = \frac{1}{2\pi}\int_a^b e^{-x^2/2}dx, \text{ where } a = \frac{c - np}{\sqrt{npq}}, \ b = \frac{d - np}{\sqrt{npq}}.$$

For example, suppose we flip a coin 100 times. Each flip has $\mu = 1/2, \sigma^2 = 1/4$ The total number of times tails comes up has mean 50 and variance 25. The probability that the number of flips is between c and d is approximated by the probability that a standard normal variable falls between $a = (c - 50)/10 \times 5 = (c - 50)/50$ and $b = (d - 50)/50$, which you can check in the tables.

Example 2 (Uniform Densities)
Suppose that each X_i is uniformly distributed between 0 and 1. Then $\mu = 1/2$ and $\sigma^2 = 1/12$. So:

$$S_n^* = \frac{S_n - n/2}{\sqrt{12n}}.$$

If you sample 100 points uniformly on $[0, 1]$, their sum has mean 50 and variance 100/12. Then

$$S_{100}^* = \frac{S_{100} - 50}{5/\sqrt{3}}.$$

The probability that the sum falls between 50 and 55 is then the same as the probability that a standard normal variable falls between 0 and $b = (55 - 50)/(5/\sqrt{3}) = \sqrt{3}$. According to tables, that is about 0.458.

Example 3 (Rolling Dice)
Roll 600 fair six-sided dice. Estimate the probability that the sum of the outcomes is more than 375.
We have:

$$S_{600}^* = \frac{S_{600} - 350}{\sqrt{600 \times 35/12}} = \frac{S_{600} - 350}{\sqrt{50 \times 35}} = \frac{S_{600} - 350}{5\sqrt{70}}$$

$S_{600} = 375$ corresponds to $S_{600}^* = 5/sqrt70 \simeq 0.5976$, and the probability that a standard normal variable falls below that is, according to tables, about 0.7257.