The Central Limit Theorem (page 288)

In the textbook, the short proof of the Central Limit Theorem involves only two equations (16) and (17). The first describes $\mathrm{E}\left[e^{itY}\right]$ and the second describes the limit of $\left(\mathrm{E}\left[e^{itY/\sqrt{N}}\right]\right)^N$ as $N \to \infty$. That limit is $e^{-t^2/2}$ by a step that appears in freshman calculus (with $a = t^2/2$):

$$\left(1 - \frac{a}{N}\right)^N \text{ approaches } e^{-a} \text{ as } N \to \infty.$$

This step comes when you define e as the limit of $(1 + \frac{1}{N})^N$.

In my 18.065 lecture of April 24, 2019 I added more detail (badly needed) to this very brief presentation in the book. This online note aims to give you the key ideas in that lecture. Start with an example:

Rolling *N* dice. For $N = 1$ we know the probabilities of seeing the outputs $X = 1, 2, 3, 4, 5, 6$:

$$p_1, p_2, p_3, p_4, p_5, p_6 = \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}$$

When we roll $N = 2$ dice, the outputs are $2, 3, \ldots, 7, \ldots, 11, 12$ with probabilities $p_2, p_3, \ldots, p_7, \ldots, p_{11}, p_{12} = \frac{1}{36}, \frac{2}{36}, \cdots, \frac{6}{36}, \cdots, \frac{2}{36}, \frac{1}{36}$. For example, $1 + 2$ and $2 + 1$ are the 2 combinations (out of 36) that go into the probability p_3.

How are those two probability vectors related? The neatest answer comes from the two "characteristic functions". Here is the relation in our example. It connects the ordinary Fourier series with those two sets of probabilities as the coefficients:

$$(\mathrm{E}[e^{itY}])^2 = \left(\frac{1}{6}e^{it} + \frac{1}{6}e^{2it} + \cdots + \frac{1}{6}e^{6it}\right)^2 = \frac{1}{36}e^{2it} + \frac{2}{36}e^{3it} + \cdots + \frac{6}{36}e^{7it} + \cdots + \frac{1}{36}e^{12it}.$$

For two dice, we square the function! For N dice, the probabilities for the possible outputs $N, \ldots, 6N$ come directly from the Nth power $(\mathrm{E}[e^{itX}])^N$ of the same 6-term sum. This is the key idea—**Go from probabilities to functions of *t***. (Fourier series would often use θ instead of t.) The crucial point is that we just take Nth powers of the function, when we are combining N independent identical trials.

Note: This valuable fact is just the Convolution Rule. The probabilities for $N = 2$ came from **convolution** of the probabilities for $N = 1$. To convolve coefficients we multiply the Fourier series:

$$\textbf{Convolve} \qquad \left(\frac{1}{6}, \ldots, \frac{1}{6}\right) * \left(\frac{1}{6}, \ldots, \frac{1}{6}\right) = \left(\frac{1}{36}, \frac{2}{36}, \cdots, \frac{2}{36}, \frac{1}{36}\right)$$

$$\textbf{Multiply} \qquad \left(\frac{1}{6}e^{it} + \cdots + \frac{1}{6}e^{6it}\right)^2 = \frac{1}{36}e^{2it} + \frac{2}{36}e^{2it} + \cdots + \frac{2}{36}e^{11it} + \frac{1}{36}e^{12it}$$

Before the step with $N \to \infty$, we have to change from X to $Y = (X - m)/\sigma$. That gives us $e^{itY} \approx 1 - \frac{1}{2}t^2$ in equation (16). Now we can work with $N \to \infty$ in equation (17).

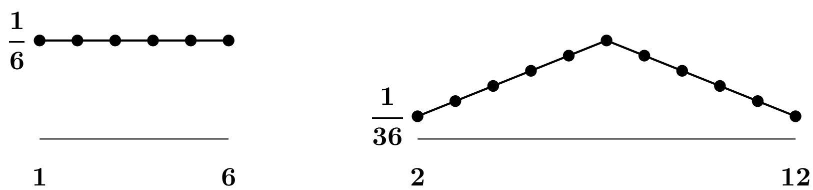
One more note about the **"Fourier transform"** that produces the series from the coefficients. At the end of the proof, we have to transform back (**inverse transform**). That step is beautiful in this problem, but we didn't show it. The Nth power of the series leads us as $N \to \infty$ to the function $e^{-t^2/2}$. This is the Fourier transform of our answer. The inverse Fourier transform (to find probabilities) will produce the limiting probability distribution for N dice or N flipped coins or N repetitions of any random trial with mean $m = 0$ and variance $\sigma^2 = 1$.

So we need the inverse transform of $e^{-t^2/2}$. The amazing answer is: **It happens to be the same function**, apart from a constant factor $1/\sqrt{2\pi}$. The limiting probability distribution for N repetitions is the **standard normal distribution**:

$$\textbf{CLT} \qquad \text{The probabilities for } Z_N = (Y_1 + \cdots + Y_N)/\sqrt{N} \text{ approach } p_{\text{normal}}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Always the Y's are the original outputs X after standardizing to $(X - m)/\sigma$. For the dice, the mean $m = 3.5$ is halfway from 1 to 6. The variance is $\sigma^2 = \frac{1}{6}\big(1 - 3.5\big)^2 + \cdots + \frac{1}{6}\big(6 - 3.5\big)^2$. It is Y that gives the standard normal distribution from the Central Limit Theorem.

Possibly you can see the bell-shaped curve $e^{-x^2/2}$ beginning to emerge from this picture for dice ($N = 1$ and $N = 2$):



For the reader's convenience we copy here those paragraphs from page 288—and also the new problem at the end of this Section V.3.

The Central Limit Theorem

In a few lines, we can justify the great limit theorem of probability. It concerns the standardized averages $Z_n = \sum (X_k - m)/\sigma\sqrt{N}$ of N independent samples $X_1, \ldots, X_N$ with mean m and variance σ^2. The central limit theorem says: **The distribution of Z_n approaches the standard normal distribution (mean zero, variance 1) as $N \to \infty$.**

The proof uses the characteristic function of the standardized variable $Y = (X - m)/\sigma$:

$$\mathrm{E}\left[e^{itY}\right] = \mathrm{E}\left[1 + itY - \frac{1}{2}t^2Y^2 + O(t^3)\right] = \mathbf{1} + \mathbf{0} - \frac{\mathbf{1}}{\mathbf{2}}\boldsymbol{t^2} + O(t^3) \tag{16}$$

Certainly $Z_N = (Y_1 + Y_2 + \cdots + Y_N)/\sqrt{N}$. So its characteristic function is a product of N identical functions of $Y/\sqrt{N}$; t is fixed. The Problem Set now ends with an example.

$$\left(\mathrm{E}\left[e^{itY/\sqrt{N}}\right]\right)^{\boldsymbol{N}} = \left[1 - \frac{1}{2}\left(\frac{t}{\sqrt{N}}\right)^2 + O\left(\frac{t}{\sqrt{N}}\right)^3\right]^{\boldsymbol{N}} \to \boldsymbol{e^{-t^2/2}} \quad \text{as} \quad N \to \infty. \tag{17}$$

That limit $\boldsymbol{e^{-t^2/2}}$ is the characteristic function of a standard normal distribution $\mathrm{N}(0, 1)$.

New Problem 9

Find m, σ^2, and $\mathbf{E}[e^{itX}] = \cos t$ if $X = 1$ with $p_1 = \frac{1}{2}$ and $X = -1$ with $p_{-1} = \frac{1}{2}$.

What are p_2, p_0, p_{-2} for the sum Z of two samples of X? Check $\mathbf{E}\left[e^{itZ}\right] = (\mathbf{E}[e^{itX}])^2$.

Using $\left(1 - \frac{C}{N}\right)^N \to e^{-C}$ from calculus, explain (17): $\left[\mathbf{E}\left(e^{itX/\sqrt{N}}\right)\right]^N \to e^{-t^2/2}$.