Solution prepared by Tobin Fricke
Prof. S.G. Rajeev, Univ. of Rochester

13. $\xi_1, \cdots \xi_n$ *are a set of n independent identically distributed random variables uniformly distributed in the interval* $[0, 1]$. *Let* μ *and* σ *be the mean and standard deviation of one of them. Define* $\zeta_n = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{n} (\xi_i - \mu)$.

When considering sums of random variables, it's almost always easiest to use the characteristic function. This is actually a result of the convolution theorem, which says that the fourier transform turns convolution into multiplication and vice versa. Pretend we have independent discrete random variables x and y and we're interested in their sum $z = x + y$. Then it's pretty easy to break that down into a sum over all possible partitions of z:

$$p_z(k) = P[z = k] = \sum_{\tau=-\infty}^{\infty} P[x = \tau] \cdot P[y = (k - \tau)] = (p_x * p_y)(k) \tag{1}$$

That has the structure of a convolution. Go to continuous random variable and a similar relation holds; the sum becomes an integral. Take the fourier transform and the probability densities turn into characteristic functions and the convolution turns into multiplication:

$$\begin{aligned} p_z(k) &= (p_x * p_y)(k) \\ F_z(t) &= (F_x \cdot F_y)(k) = F_x(k) F_y(k) \end{aligned} \tag{2}$$

Then of course we may take the inverse transform to get back to a probability density, from the characteristic function, of the sum.

(This tour through characteristic functions is not only labor-saving in the analysis, but is labor-saving in numerical computations, too. It turns out that it takes time proportional to $n \log n$ to compute a fourier transform of n samples whereas it takes time proportional to n^2 to compute a convolution in the "naïve manner". For this reason convolutions are almost always computed using the relation $x * y = F^{-1}\{F\{x\} \cdot F\{y\}\}$.)

(a) *What is the probability density function of* ξ_n *for* $n \in \{1, 2, 3\}$*?*

For the uniform distribution over $[0, 1]$, we know the distribution's mean ($\mu = 1/2$) and variance ($\sigma = \sqrt{1/12}$). It's not clear here whether the problem is looking for the *sample mean and variance* which are estimated from an ensemble of samples of the random variable, or the *distribution* mean and variance, which are known theoretically from the distribution. I'll assume the latter. The former converges to the latter given many samples anyway.

The characteristic function is defined as:

$$f_\eta = \langle e^{it\eta} \rangle = \int_{-\infty}^{\infty} e^{itx} p_\eta(x) dx \tag{3}$$

Instead of the random variables ξ_i distributed uniformly over $[0, 1]$, consider the transformed random variables $\xi_i' = (\xi_i - \mu)/\sigma$; from here we'll consider these transformed variables instead. These new variables are also unformly distributed, now over the interval $\left[-\frac{1}{2\sigma}, \frac{1}{2\sigma}\right]$ with density σ. The characteristic function for these guys is:

$$f_{\xi_i'}(t) = \int_{-1/2\sigma}^{1/2\sigma} \sigma e^{itx} dx = \frac{2\sigma}{t} \sin\left(\frac{t}{2\sigma}\right) = \operatorname{sinc}\left(\frac{t}{2\sigma}\right)$$

We immediately have the answer to the $n = 1$ case:

$$\boxed{p_{\zeta_1}(x) = \begin{cases} \sigma & -\frac{1}{2\sigma} \le x \le \frac{1}{2\sigma} \\ 0 & \text{otherwise} \end{cases}}$$

For $n = 2$ the problem is simple enough that we don't actually have to do the whole jaunt through fourier space (via the characteristic function). I call the function whose value is constant (in this case, with value σ) in some finite interval (in this case $\left[-\frac{1}{2\sigma}, \frac{1}{2\sigma}\right]$), and zero elsewhere a square "gate" function. We have two identical gate functions sitting on the x axis. If we visualize the convolution process as a "sliding dot product" (or "delay-and-sum") operation, we realize that the convolution of these two gates will be a triangle function, also centered on the origin. The triangle function will have twice the horizontal extent as the original gate functions, so it will be nonzero in the region $\left[-\frac{1}{\sigma}, \frac{1}{\sigma}\right]$. Moreover, since this is a probability density, it must integrate to one; therefore the height of the triangle is σ and we may write down the answer.

$$\boxed{p_{\zeta_2}(x) = \begin{cases} \sigma - \frac{|x|}{\sigma} & -\frac{1}{\sigma} \le x \le \frac{1}{\sigma} \\ 0 & \text{otherwise} \end{cases}}$$

For $n = 3$ and higher, it becomes simpler just to apply what we know about characteristic functions. We recall $f_{a\eta+b}(t) = e^{itb} f_\eta(at)$ and compute:

$$f_{\zeta_3}(t) = f_{\frac{1}{\sqrt{3}}(\xi'_1+\xi'_2+\xi'_3)}(t) = \left(f_{\xi'}\left(\frac{1}{\sqrt{3}}t\right)\right)^3 = \frac{24\sigma^3\sqrt{3}}{t^3}\sin^3\left(\frac{t}{2\sigma\sqrt{3}}\right)$$

We can take inverse fourier transform to get a probability density:

$$p_\eta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f_\eta(t)e^{-itx}dt \tag{4}$$

(b) *What is the probability density function of ζ_n as $n \to \infty$?*
Consider what happens to the characteristic function:

$$f_{\zeta_\infty}(t) = \lim_{n\to\infty} f_{\zeta_n}(t) = \lim_{n\to\infty}\left[f_{\xi'}\left(\frac{t}{\sqrt{n}}\right)\right]^n = \lim_{n\to\infty}\left(\frac{2\sigma\sqrt{n}}{t}\sin\frac{t}{2\sigma\sqrt{n}}\right)^n$$

This converges to a Gaussian. One way to see this is to expand $f_{\xi'}$ in a power series. Also, this is as good a time as any to remember that $\sigma = \sqrt{1/12}$. We use the identity $\lim_{n\to\infty}(1 + x/n)^n = \exp x$ which was derived in the solutions to the first problem set.

$$f_{\zeta_\infty}(t) = \lim_{n\to\infty}\left[1 - \frac{t^2}{2n} + O(n^{-2})\right]^n = \exp\left\{-t^2/2\right\}$$

As any electrical engineer will tell you, the fourier transform of a Gaussian is a Gaussian:

$$p_{\zeta_\infty}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f_{\zeta_\infty}(t)e^{-ikt}dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} f_{\zeta_\infty}(t)e^{-t^2/2}e^{-ikt}dt = \frac{1}{2\pi}\exp\{-x^2/2\}$$

(Though the normalization is not correct here. Renormalizing we find $p(x) = \sqrt{\frac{2}{\pi}}e^{-x^2/2}$.)
This is the result we should have expected from the central limit theorem.

14. Write a computer program that generates n uniform random variables, and calculates ξ_n as defined above. Then repeat this a large number N times to get a simulated data sample for ξ_n and compute the relative frequencies in bins of equal width. For $N = 1000$ and $n \in \{1, 2, 3, 100, 1000\}$, compare the simulation with the theoretical predictions above, by plotting or preparing a table.

    Coming soon!

15. Let ξ_1, ξ_2 be independent identically distributed random variables with the Cauchy (or Lorentzian) probability distribution function ($p(x) = (1/\pi)/(1 + x^2)$).

   (a) What is the probability density of the sum $\xi_1 + \xi_2$?
   The Cauchy distribution has characteristic function:

$$f_\xi(t) = \exp\{-|t|\} \tag{5}$$

   The sum $\xi_1 + \xi_2$ therefore has characteristic function:

$$f_{\xi_1+\xi_2}(t) = \exp\{-|t|\} \exp\{-|t|\} = \exp\{-2|t|\}$$

   The inverse transform is:

$$p_{\xi_1+\xi_2}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} e^{-2|t|} dt$$

   In general we find (using Jordan's lemma):

$$p_{\xi_1+\cdots+\xi_n} = \frac{n}{\pi(n^2 + x^2)}$$

   In particular:

$$p_{\xi_1+\xi_2} = \frac{2}{\pi(4 + x^2)}$$

   (b) *If there are n such independent identially distributed Cauchy variables $\xi_1, \cdots \xi_n$, what is the probability density function of their average in the limit $n \to \infty$?*
   You will find that the distribution of the mean of an ensemble containing many Cauchy-distributed random variables again is distributed according to the Cauchy distribution. The central limit theorem does not apply here, because it assumes finite variance!
   See `http://en.wikipedia.org/wiki/Cauchy_distribution`.