The central limit theorem.

To prove the central limit theorem we make use of the Fourier transform which is one of the most useful tools in pure and applied analysis and is therefore interesting in its own right.

We say a $f : \mathbf{R} \to \mathbf{C}$ is **summable** if

$$\int |f(x)|\, dx \;<\; \infty.$$

For any such function we define its **Fourier transform**

$$\hat{f} : \mathbf{R} \to \mathbf{C}$$

by setting

$$\hat{f}(t) \;=\; \int e^{-itx}\, f(x)\, dx \qquad \text{for } t \in \mathbf{R}.$$

Note that $f \mapsto \hat{f}$ is linear.

For $a \in \mathbf{R}$ set $\tau_a(x) = x + a$. Then

$$\widehat{f \circ \tau_a}(t) \;=\; e^{ita}\, \hat{f}(t).$$

Indeed,

$$\begin{aligned}
\widehat{f \circ \tau_a}(t) \;&=\; \int e^{-itx}\, f(x+a)\, dx \\
&=\; \int e^{-it(y-a)}\, f(y)\, dy \qquad \text{substitute } y - a \text{ for } x \\
&=\; e^{ita} \int e^{-ity}\, f(y)\, dy \\
&=\; e^{ita}\, \hat{f}(t).
\end{aligned}$$

As a corollary we obtain that if f' is summable then

$$\widehat{f'}(t) \;=\; \lim_{h \to 0} \frac{\widehat{f \circ \tau_h} \;-\; f}{h}(t) \;=\; \lim_{h \to 0} \frac{e^{ith} \;-\; 1}{h}\, \hat{f}(t) \;=\; it\, \hat{f}(t).$$

For $c > 0$ set $\delta_c(x) = cx$. Then

$$\widehat{f \circ \delta_c} \;=\; \frac{1}{c}\, \hat{f} \circ \delta_{\frac{1}{c}}.$$

Indeed,

$$\begin{aligned}
\widehat{f \circ \delta_c}(t) \;&=\; \int e^{-itx}\, f(cx)\, dx \\
&=\; \int e^{-it\frac{y}{c}}\, f(y)\, \frac{dy}{c} \qquad \text{substitute } \tfrac{y}{c} \text{ for } x \\
&=\; \frac{1}{c}\, \hat{f} \circ \delta_{\frac{1}{c}}(t).
\end{aligned}$$

If f and g are summable then so is $f * g$ and

$$\widehat{f \,*\, g} \;=\; \hat{f}\, \hat{g}.$$

Indeed,

$$\begin{aligned}
\widehat{f * g}(t) &= \int e^{-itx} \int f(x - y)\, g(y)\, dy\, dx \\
&= \int \int e^{-it(x-y)} f(x - y)\, e^{-ity} g(y)\, dx\, dy \\
&= \int \int e^{-itw} f(x)\, dw\, e^{-ity} g(y)\, dy \qquad \text{for each } y \text{ substitute } w + y \text{ for } x \\
&= \hat{f}(t)\, \hat{g}(t).
\end{aligned}$$

The Fourier inversion formula. Suppose f is twice differentiable and each of f,f' and f'' are summable. Then

$$f(x) = \frac{1}{2\pi} \int e^{itx} \hat{f}(t)\, dt \qquad \text{for each } t \in \mathbf{R}.$$

Remark. You don't need to assume this much, but the proof gets more technical.

Proof. It will suffice to prove that

$$f(0) = \frac{1}{2\pi} \int \hat{f}(t)\, dt \tag{1}$$

for all f satisfying the hypotheses because if this is so and $x \in \mathbf{R}$ we may set $g = f \circ \tau_x$ and observed that

$$f(x) = g(0) = \int \hat{g}(t)\, dt = \frac{1}{2\pi} \int e^{itx} \hat{f}(t)\, dt.$$

To prove (1) we will make use of the

Lemma.

$$\lim_{R \uparrow \infty} \int_0^R \frac{\sin x}{x}\, dx = \frac{\pi}{2}.$$

Proof. Note that

$$e^{-tx} \sin x = \frac{d}{dx} \frac{e^{-tx}}{t^2 + 1} (\cos x + t \sin x).$$

We have

$$\begin{aligned}
\int_0^R \frac{\sin x}{x}\, dx &= \int_0^R \sin x \int_0^\infty e^{-tx} dt\, dx \\
&= \int_0^\infty \int_0^R e^{-tx} \sin x\, dx\, dt \\
&= \int_0^\infty \frac{1}{t^2 + 1} (1 - e^{-tR} (\cos R + t \sin R))\, dt \\
&\to \frac{\pi}{2} \qquad \text{as } R \uparrow \infty.
\end{aligned}$$

$\square$

We have $f(x) = f(0) + xq(x)$ where we have set $q(x) = \int_0^1 f(\lambda x) d\lambda$. Let

$$I(R) = \int \frac{\sin R\, x}{x} f(x)\, dx.$$

Then

$$\int \hat{f}(t)\,dt \;=\; \lim_{R\to\infty}\int_{-R}^{R}\int e^{-itx} f(x)\,dx\,dt \;=\; \int\int_{-R}^{R} e^{-itx}\,dt\,f(x)\,dx \;=\; I(R).$$

Now

$$I(R) \;=\; I_1(R,\epsilon) \;+\; I_2(R,\epsilon) \;+\; I_3(R,\epsilon)$$

for any $\epsilon > 0$ where we have set

$$\begin{aligned} I_1(R,\epsilon) \;&=\; \int_{|x|>\epsilon} \sin Rx\,\frac{f(x)}{x}\,dx, \\ I_2(R,\epsilon) \;&=\; f(0)\int_{|x|\le\epsilon} \frac{\sin Rx}{x}\,dx \text{ and} \\ I_3(R,\epsilon) \;&=\; \int_{|x|\le\epsilon} \sin Rx\,q(x)\,dx. \end{aligned}$$

Approximating $x \mapsto \frac{f(x)}{x}$ by step functions on $|x| > \epsilon$ we find that $\lim_{R\uparrow\infty} I_1(R,\epsilon) = 0$ for each $\epsilon > 0$. Substituing $\frac{w}{R}$ for x we find that

$$I_2(R,\epsilon) \;=\; 2\,f(0)\int_0^{R\epsilon}\frac{\sin w}{w}\,dw.$$

Moreover,

$$|I_3(R,\epsilon) \;\le\; 2\,\epsilon\,\sup\{|f'(x)| : |x| \le \epsilon\}.$$

It should now be clear that by making R large we can make $I(R)$ as close as we like to

$$2\,f(0)\,\lim_{R\uparrow\infty}\int_0^R \frac{\sin w}{w}\,dw \;=\; \pi\,f(0).$$

The Gaussian. Set

$$g(x) \;=\; \frac{1}{I}\,e^{\frac{-x^2}{2}} \qquad \text{for } x \in \mathbf{R}$$

where we have set

$$I \;=\; \int e^{\frac{-x^2}{2}}\,dx.$$

This function is called the **Gaussian**; let us compute its Fourier transform. We have

$$\frac{d}{dt}\,\hat{g}(t) \;=\; \frac{i}{I}\int e^{-itx}\,d(e^{\frac{-x^2}{2}}) \;=\; \frac{t}{I}\int e^{-itx}\,e^{\frac{-x^2}{2}}\,dx \;=\; -t\,\hat{g}(t)$$

from which we infer that

$$\hat{g}(t) \;=\; \hat{g}(0)\,e^{\frac{-t^2}{2}} \;=\; I\,g(t).$$

From the Fourier inversion formula we obtain

$$\frac{1}{I} \;=\; g(0) \;=\; \frac{1}{2\pi}\int \hat{g}(t)\,dt \;=\; \frac{I}{2\pi},$$

which implies that

$$I \;=\; \sqrt{2\pi}$$

and that

$$\hat{g}(t) \;=\; \sqrt{2\pi}\,g(t).$$

The characteristic function of a random variable. Suppose X is a random variable for which $E(|X|) < \infty$. We set

$$\chi_X(t) \,=\, E(e^{itX}) \qquad \text{for } t \in \mathbf{R}$$

and call this function the **characteristic function of** X. If X is continuous we have

$$\chi_X(t) \,=\, \int e^{itx}\, f_X(x)\, dx \,=\, \widehat{f_X}(-t).$$

If X is discrete, the characteristic function has a simple Fourier analytic interpretation, but we shall not go into that.

Now suppose that

$$E(|X|^j) \,<\, \infty \qquad \text{for } j = 1, 2, 3.$$

From Taylor's theorem we obtain for any $t \neq 0$ that

$$\chi_X(t) \,=\, \chi_X(0) \,+\, \chi_X'(0)\, t \,+\, \chi_X''(0)\, \frac{t^2}{2} \,+\, \chi_X'''(s)\, \frac{t^3}{6}$$

for some $s \in (-|t|, |t|)$. This gives

$$|\chi_X(t) \,-\, (1 \,+\, t\, E(X) \,+\, \frac{t^2}{2}\, E(X^2))| \,\leq\, \frac{|t|^3}{6}\, E(|X|^3).$$

Suppose $c > 0$. We have

$$f_{cX}(y) \,=\, \lim_{h \downarrow 0} \frac{P(y \,-\, h \,<\, cX \,<\, y + h)}{2\, h} \,=\, \lim_{h \downarrow 0} \frac{1}{c} \frac{P(\frac{y-h}{c} \,<\, cX \,<\, \frac{y+h}{c})}{\frac{2\,h}{c}} \,=\, \frac{1}{c}\, f_X(\frac{y}{c}).$$

This implies that

$$\chi_{cX}(t) \,=\, \widehat{f_{cX}}(-t) \,=\, \frac{1}{c} \widehat{f_X \circ \delta_{\frac{1}{c}}}(-t) \,=\, \frac{1}{c}\, c\, \widehat{f_X} \circ \delta_c(-t) \,=\, \widehat{f_X}(-ct) \,=\, \chi_X(ct).$$

The proof of the central limit theorem. Let $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of independent identically distributed random variables, each having mean 0 and variance 1. We will leave it to the reader to spell out the hypotheses on the X_i's which will allow us to make use of the foregoing. Let χ be the common characteristic function of these variables. For each positive integer n let $S_n = \sum_{i=1}^n X_i$. We have

$$\chi_{\frac{S_n}{\sqrt{n}}}(t) \,=\, \chi_{S_n}(\frac{t}{\sqrt{n}} \,=\, \chi(\frac{t}{\sqrt{n}})^n \,=\, (1 \,-\, \frac{t^2}{2\, n} \,+\, \theta(t)\, \frac{t^3}{6\, n^{3/2}}\, E(|X|^3))^n \,\rightarrow\, e^{\frac{t^2}{2}}$$

where $|\theta(t)| \leq 1$ for each t. Thus if $-\infty < a < b < \infty$ we have

$$\begin{aligned}
P(a \,<\, \frac{S_n}{\sqrt{n}} \,<\, b) \,&=\, \int_a^b f_{\frac{S_n}{\sqrt{n}}}(x)\, dx \\
&=\, \frac{1}{2\pi} \int_a^b \int e^{itx} \widehat{f_{\frac{S_n}{\sqrt{n}}}}(t)\, dt\, dx \\
&=\, \frac{1}{2\pi} \int_a^b \int e^{itx} \chi(\frac{-t}{\sqrt{n}})^n\, dt\, dx \\
&\rightarrow\, \frac{1}{2\pi} \int_a^b \int e^{itx} e^{\frac{-t^2}{2}}\, dt\, dx \text{ as } n \rightarrow \infty \\
&=\, \frac{1}{\sqrt{2\pi}} \int_a^b \hat{g}(-x)\, dx \\
&=\, \int_a^b g(-x)\, dx \\
&=\, \int_a^b g(x)\, dx.
\end{aligned}$$

$\square$