A Direct Proof of the Central Limit Theorem

Shoichi Midorikawa

Let $X_1,\ X_2,\ \cdots,\ X_n$ be independent and identically distributed (i.i.d.) random variables with common probability density function $f(X)$. For simplicity, assume that $f(X)$ has mean 0 and variance σ^2. That is,

$$\int x f(x)\,dx = 0 \tag{1}$$

$$\int x^2 f(x)\,dx = \sigma^2 \tag{2}$$

Let $f_{\bar{X}}(\bar{x})$ denote the probability density function of the random variable $\bar{X} = (X_1 + X_2 + \cdots + X_n)/\sqrt{n}\sigma$. Then

$$f_{\bar{X}}(\bar{x}) = \int \delta\left(\bar{x} - \frac{x_1 + x_2 + \cdots + x_n}{\sqrt{n}\sigma}\right) f(x_1) f(x_2) \cdots f(x_n)\,dx_1 dx_2 \cdots dx_n \tag{3}$$

is obtained. Using the integral representation of the delta function,

$$\delta\left(\bar{x} - \frac{x_1 + x_2 + \cdots + x_n}{\sqrt{n}\sigma}\right) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ik\left(\bar{x} - (x_1 + x_2 + \cdots + x_n)/\sqrt{n}\sigma\right)}\,dk$$

and substituting this into (3), we obtain

$$\begin{aligned} f_{\bar{X}}(\bar{x}) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik\bar{x}} \prod_{i=1}^{n} \int e^{-ikx_i/\sqrt{n}\sigma} f(x_i)\,dx_i \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik\bar{x}} \left[\int e^{-ikx/\sqrt{n}\sigma} f(x)\,dx\right]^n \end{aligned} \tag{4}$$

The function inside the brackets [ ] in (4),

$$\varphi\left(\frac{k}{\sqrt{n}\sigma}\right) = \int e^{-ikx/\sqrt{n}\sigma} f(x)\,dx \tag{5}$$

is the characteristic function of the distribution $f(X)$. This differs slightly from the usual definition in sign convention, but this is not essential. To match the standard definition exactly, simply replace k by $-k$ in (4).

Now expand $e^{-ikx/\sqrt{n}\sigma}$ by Taylor's formula in a form with remainder:

$$\begin{aligned} e^{-ikx/\sqrt{n}\sigma} &= 1 + i\frac{k}{\sqrt{n}\sigma}\int_0^x (x-t)' e^{-ikt/\sqrt{n}\sigma}\,dt \\ &= 1 + i\frac{k}{\sqrt{n}\sigma}\left\{\Big[(x-t)e^{-ikt/\sqrt{n}\sigma}\Big]_0^x + i\frac{k}{\sqrt{n}\sigma}\int_0^x (x-t) e^{-ikt/\sqrt{n}\sigma}\,dt\right\} \\ &= 1 - i\frac{k}{\sqrt{n}\sigma}\,x - \left(i\frac{k}{\sqrt{n}\sigma}\right)^2 \int_0^x \left[\frac{(x-t)^2}{2!}\right]' e^{-ikt/\sqrt{n}\sigma}\,dt \\ &= 1 - i\frac{k}{\sqrt{n}\sigma}\,x - \left(i\frac{k}{\sqrt{n}\sigma}\right)^2 \left\{\left[\frac{(x-t)^2}{2!}e^{-ikt/\sqrt{n}\sigma}\right]_0^x + i\frac{k}{\sqrt{n}\sigma}\int_0^x \frac{(x-t)^2}{2!} e^{-ikt/\sqrt{n}\sigma}\,dt\right\} \\ &= 1 - i\frac{k}{\sqrt{n}\sigma}\,x - \frac{k^2}{2n\sigma^2}\,x^2 + i\frac{k^3}{n^{3/2}\sigma^3}\int_0^x \frac{(x-t)^2}{2!} e^{-ikt/\sqrt{n}\sigma}\,dt \end{aligned}$$

Define

$$R(x) = \int_0^x \frac{(x-t)^2}{2!} e^{-ikt/\sqrt{n}\sigma}\, dt \tag{6}$$

Then we can write

$$e^{-ikx/\sqrt{n}\sigma} = 1 - i\frac{k}{\sqrt{n}\sigma}\, x - \frac{k^2}{2n\sigma^2}\, x^2 + i\frac{k^3}{n^{3/2}\sigma^3} R(x) \tag{7}$$

Moreover, it is easy to see that

$$|R(x)| \le \int_0^{|x|} \frac{(x-t)^2}{2!}\, dt = \frac{|x|^3}{3!}$$

To evaluate (**??**), use (**??**) together with $\int f(x)\, dx = 1$, and also (1) and (2). Then

$$\begin{aligned} \int e^{-ikx/\sqrt{n}\sigma} f(x)\, dx &= \int \left\{ 1 - i\frac{k}{\sqrt{n}\sigma}\, x - \frac{k^2}{2n\sigma^2}\, x^2 + i\frac{k^3}{n^{3/2}\sigma^3} R(x) \right\} f(x)\, dx \\ &= 1 - i\frac{k}{\sqrt{n}\sigma} \int x f(x)\, dx - \frac{k^2}{2n\sigma^2} \int x^2 f(x)\, dx + i\frac{k^3}{n^{3/2}\sigma^3} \int R(x) f(x)\, dx \\ &= 1 - \frac{k^2}{2n} + i\frac{k^3}{n^{3/2}\sigma^3} \int R(x) f(x)\, dx \end{aligned} \tag{8}$$

Thus, if we write $C = \int R(x) f(x)\, dx$, then

$$|C| \le \int |R(x)| f(x)\, dx = \frac{1}{3!} \int |x|^3 f(x)\, dx$$

We now assume that the third absolute moment exists, namely, $\int |x|^3 f(x)\, dx < \infty$. In that case, clearly $|C| < \infty$. Therefore,

$$\lim_{n\to\infty} \left[\int e^{-ikx/\sqrt{n}\sigma} f(x)\, dx \right]^n = \lim_{n\to\infty} \left[1 - \frac{k^2}{2n} + i\frac{k^3}{n^{3/2}\sigma^3}\, C \right]^n$$

can be evaluated.

Set

$$U = \frac{k^2}{2n} - i\frac{k^3}{n^{3/2}\sigma^3}\, C$$

Then

$$\left[1 - \frac{k^2}{2n} + i\frac{k^3}{n^{3/2}\sigma^3}\, C \right]^n = \Big[1 - U \Big]^n = e^{n \ln(1-U)}$$

Now,

$$\ln(1-U) = -U - \frac{U^2}{2} - \frac{U^3}{3} - \frac{U^4}{4} - \cdots \qquad (|U| < 1)$$

so that

$$|\ln(1-U) + U| = \left| \frac{U^2}{2} + \frac{U^3}{3} + \frac{U^4}{4} + \cdots \right| \le \frac{|U|^2}{2} + \frac{|U|^3}{2} + \frac{|U|^4}{2} + \cdots \le \frac{1}{2}\, \frac{|U|^2}{1 - |U|}$$

Therefore, if $|U| \leq 1/2$, then $1-|U| \geq 1/2$, and hence

$$|\ln(1-U)+U| \leq |U|^2$$

Thus,

$$\lim_{n\to\infty} |n\ln(1-U)+nU| \leq \lim_{n\to\infty} n|U|^2 = 0$$

and therefore

$$\lim_{n\to\infty} n\ln(1-U) = -\lim_{n\to\infty} nU = -\frac{k^2}{2}$$

Hence,

$$\lim_{n\to\infty}\left[\int e^{-ikx/\sqrt{n}\sigma} f(x)\,dx\right]^n = e^{-k^2/2} \tag{9}$$

is obtained. Substituting the right-hand side of (**??**) into (4), we find

$$\begin{aligned}\lim_{n\to\infty} f_{\bar{X}}(\bar{x}) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\ \exp\left[ik\bar{x}-\frac{k^2}{2}\right] \\ &= \frac{1}{2\pi}e^{-\bar{x}^2/2}\int_{-\infty}^{\infty} dk \exp\left[-\frac{1}{2}\left(k-i\bar{x}\right)^2\right]\end{aligned} \tag{10}$$

Since

$$\int_{-\infty}^{\infty} dk \exp\left[-\frac{1}{2}\left(k-i\bar{x}\right)^2\right] = \int_{-\infty}^{\infty} dk' e^{-k'^2/2} = \sqrt{2\pi}$$

we conclude from (**??**) that

$$\lim_{n\to\infty} f_{\bar{X}}(\bar{x}) = \frac{1}{\sqrt{2\pi}}e^{-\bar{x}^2/2}$$

That is, the limiting distribution is the standard normal distribution $N(0,\ 1)$, with mean 0 and variance 1.

More generally, if the random variables $X_1,\ X_2,\ \cdots,\ X_n$ are independent and identically distributed with mean μ and variance σ^2, then the random variable $\bar{X} = (X_1+X_2+\cdots+X_n-n\mu)/\sqrt{n}\sigma$ is approximately distributed as the standard normal distribution $N(0,\ 1)$ when n is sufficiently large, regardless of the detailed shape of the original probability density function.

Example 1 Standard normal distribution $f(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$

$$\begin{aligned} f_{\bar{X}}(\bar{x}) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik\bar{x}}\left[\frac{1}{\sqrt{2\pi}}\int e^{-ikx/\sqrt{n}-x^2/2}\,dx\right]^n \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik\bar{x}}\left[\frac{e^{-k^2/2n}}{\sqrt{2\pi}}\int e^{-(x+ik/\sqrt{n})^2/2}\,dx\right]^n \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk e^{-k^2/2+ik\bar{x}} \\ &= \frac{e^{-\bar{x}^2/2}}{2\pi}\int_{-\infty}^{\infty} dk e^{-(k-i\bar{x})^2/2} \\ &= \frac{1}{\sqrt{2\pi}}e^{-\bar{x}^2/2}\end{aligned}$$

Thus the random variable $\bar{X} = (X_1 + X_2 + \cdots + X_n)/\sqrt{n}$ follows the standard normal distribution $N(0,\ 1)$ exactly.

Example 2 Uniform distribution $f(x) = \begin{cases} 1 & \text{for } -\frac{1}{2} \le x \le \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$

In this case, $\sigma^2 = \displaystyle\int_{-1/2}^{1/2} x^2\, dx = \frac{1}{12}$ so that
$\bar{X} = (X_1 + X_2 + \cdots + X_n)/\sqrt{n}\sigma = \sqrt{12/n}\,(X_1 + X_2 + \cdots + X_n)$
and its distribution is

$$\begin{aligned}
f_{\bar{X}}(\bar{x}) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik\bar{x}} \left[\int_{-1/2}^{1/2} e^{-i\sqrt{12/n}\,kx}\, dx\right]^n \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik\bar{x}} \left(\frac{\sin\sqrt{\frac{3}{n}}\,k}{\sqrt{\frac{3}{n}}\,k}\right)^n \\
&\approx \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik\bar{x}} \left[1 - \frac{k^2}{2n}\right]^n \\
&\approx \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{-k^2/2 + ik\bar{x}} \\
&= \frac{1}{\sqrt{2\pi}} e^{-\bar{x}^2/2}
\end{aligned}$$

That is, it is approximately the normal distribution $N(0,\ 1)$.

Example 3 Exponential distribution $f(x) = \lambda e^{-\lambda x}$ $(x \ge 0)$

$$\begin{aligned}
\mu &= \lambda\int_0^{\infty} x e^{-\lambda x}\, dx = \frac{1}{\lambda} \\
\sigma^2 &= \lambda\int_0^{\infty} x^2 e^{-\lambda x}\, dx - \mu^2 = \frac{1}{\lambda^2}
\end{aligned}$$

Hence
$\bar{X} = (X_1 + X_2 + \cdots + X_n)/\sqrt{n}\sigma = \lambda(X_1 + X_2 + \cdots + X_n)/\sqrt{n}$
has distribution

$$\begin{aligned}
f_{\bar{X}}(\bar{x}) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik\bar{x}} \left[\lambda\int_0^{\infty} e^{-\lambda(1+ik/\sqrt{n})x}\, dx\right]^n \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik\bar{x}} \left[\frac{1}{1 + i\frac{k}{\sqrt{n}}}\right]^n
\end{aligned}$$

Now,

$$\left[\frac{1}{1 + i\frac{k}{\sqrt{n}}}\right]^n = \left[\frac{1 - i\frac{k}{\sqrt{n}}}{1 + \frac{k^2}{n}}\right]^n \approx \left[1 - i\frac{k}{\sqrt{n}} - \frac{k^2}{n}\right]^n \approx e^{-i\sqrt{n}k} e^{-k^2/2}$$

so that

$$f_{\bar{X}}(\bar{x}) \approx \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\bar{x} - \sqrt{n}\right)^2\right]$$

Thus it is approximately the normal distribution $N(\sqrt{n},\ 1)$.