Central Limit Theorem

Let $X_1, X_2, \ldots, X_N$ be a set of N independent random variates and each X_i have an *arbitrary* probability distribution $P(x_1, \ldots, x_N)$ with mean μ_i and a finite variance σ_i^2. Then the normal form variate

$$X_{\mathrm{norm}} \equiv \frac{\sum_{i=1}^{N} x_i - \sum_{i=1}^{N} \mu_i}{\sqrt{\sum_{i=1}^{N} \sigma_i^2}} \tag{1}$$

has a limiting cumulative distribution function which approaches a normal distribution.

Under additional conditions on the distribution of the addend, the probability density itself is also normal (Feller 1971) with mean $\mu = 0$ and variance $\sigma^2 = 1$. If conversion to normal form is not performed, then the variate

$$X \equiv \frac{1}{N} \sum_{i=1}^{N} x_i \tag{2}$$

is normally distributed with $\mu_X = \mu_x$ and $\sigma_X = \sigma_x / \sqrt{N}$.

$$
\begin{aligned}
\mathcal{F}_f^{-1}[P_X(f)](x) &\equiv \int_{-\infty}^{\infty} e^{2\pi i f X} P(X)\,dX \\
&= \int_{-\infty}^{\infty} \sum_{n=0}^{\infty} \frac{(2\pi i f X)^n}{n!} P(X)\,dX \\
&= \sum_{n=0}^{\infty} \frac{(2\pi i f)^n}{n!} \int_{-\infty}^{\infty} X^n P(X)\,dX \\
&= \sum_{n=0}^{\infty} \frac{(2\pi i f)^n}{n!} \langle X^n \rangle.
\end{aligned}
$$

Now write

$$
\begin{aligned}
\langle X^n \rangle &= \langle N^{-n} (x_1 + x_2 + \ldots + x_N)^n \rangle \\
&= \int_{-\infty}^{\infty} N^{-n} (x_1 + \ldots + x_N)^n P(x_1) \cdots P(x_N)\,dx_1 \cdots dx_N,
\end{aligned}
$$

so we have

$$\mathcal{F}_f^{-1}\left[P_X(f)\right](x) = \sum_{n=0}^{\infty} \frac{(2\pi i f)^n}{n!} \langle X^n \rangle$$

$$= \sum_{n=0}^{\infty} \frac{(2\pi i f)^n}{n!} \int_{-\infty}^{\infty} N^{-n} (x_1 + \ldots + x_N)^n \times P(x_1) \cdots P(x_N)\, d x_1 \cdots d x_N$$

$$= \int_{-\infty}^{\infty} \sum_{n=0}^{\infty} \left[\frac{2\pi i f (x_1 + \ldots + x_N)}{N}\right]^n \frac{1}{n!} P(x_1) \cdots P(x_N)\, d x_1 \cdots d x_N$$

$$= \int_{-\infty}^{\infty} e^{2\pi i f (x_1 + \ldots + x_N)/N} P(x_1) \cdots P(x_N)\, d x_1 \cdots d x_N$$

$$= \left[\int_{-\infty}^{\infty} e^{2\pi i f x_1/N} P(x_1)\, d x_1\right] \times \cdots \times \left[\int_{-\infty}^{\infty} e^{2\pi i f x_N/N} P(x_N)\, d x_N\right]$$

$$
\begin{aligned}
&= \left[\int_{-\infty}^{\infty} e^{2\pi i f x/N} P(x)\, dx \right]^N \\
&= \left\{ \int_{-\infty}^{\infty} \left[1 + \left(\frac{2\pi i f}{N} \right) x + \frac{1}{2} \left(\frac{2\pi i f}{N} \right)^2 x^2 + \ldots \right] P(x)\, dx \right\}^N \\
&= \left[1 + \frac{2\pi i f}{N} \langle x \rangle - \frac{(2\pi f)^2}{2N^2} \langle x^2 \rangle + O\left(N^{-3}\right) \right]^N \\
&= \exp\left\{ N \ln \left[1 + \frac{2\pi i f}{N} \langle x \rangle - \frac{(2\pi f)^2}{2N^2} \langle x^2 \rangle + O\left(N^{-3}\right) \right] \right\}.
\end{aligned}
$$

Now expand

$$
\ln(1+x) = x - \frac{1}{2} x^2 + \frac{1}{3} x^3 + \ldots,
$$

$$
\begin{aligned}
\mathcal{F}_f^{-1}\left[P_X(f)\right](x) &\approx \exp\left\{N\left[\frac{2\pi i f}{N}\langle x\rangle - \frac{(2\pi f)^2}{2N^2}\langle x^2\rangle + \frac{1}{2}\frac{(2\pi i f)^2}{N^2}\langle x\rangle^2 + O\left(N^{-3}\right)\right]\right\} \\
&= \exp\left[2\pi i f\langle x\rangle - \frac{(2\pi f)^2\left(\langle x^2\rangle - \langle x\rangle^2\right)}{2N} + O\left(N^{-2}\right)\right] \\
&\approx \exp\left[2\pi i f\mu_x - \frac{(2\pi f)^2\sigma_x{}^2}{2N}\right],
\end{aligned}
$$

since

$$\mu_x \equiv \langle x \rangle$$
$$\sigma_x^2 \equiv \langle x^2 \rangle - \langle x \rangle^2.$$

Taking the Fourier transform,

$$P_X \equiv \int_{-\infty}^{\infty} e^{-2\pi i f x} \mathcal{F}^{-1} [P_X(f)] \, d f$$
$$= \int_{-\infty}^{\infty} e^{2\pi i f (\mu_x - x) - (2\pi f)^2 \sigma_x^2 / 2N} \, d f.$$

This is of the form

$$\int_{-\infty}^{\infty} e^{i a f - b f^2} \, d f,$$

where $a \equiv 2\pi(\mu_x - x)$ and $b \equiv (2\pi\sigma_x)^2 / 2N$. But this is a Fourier transform of a Gaussian function, so

$$\int_{-\infty}^{\infty} e^{i a f - b f^2} \, d f = e^{-a^2/4b} \sqrt{\frac{\pi}{b}}$$

$$P_X = \sqrt{\frac{\pi}{\frac{(2\pi\sigma_x)^2}{2N}}} \exp\left\{\frac{-[2\pi(\mu_x - x)]^2}{4\frac{(2\pi\sigma_x)^2}{2N}}\right\}$$

$$= \sqrt{\frac{2\pi N}{4\pi^2\sigma_x{}^2}} \exp\left[-\frac{4\pi^2(\mu_x - x)^2 2N}{4\cdot 4\pi^2\sigma_x{}^2}\right]$$

$$= \frac{\sqrt{N}}{\sigma_x\sqrt{2\pi}} e^{-(\mu_x - x)^2 N/2\sigma_x{}^2},$$

But $\sigma_X = \sigma_x / \sqrt{N}$ and $\mu_X = \mu_x$, so

$$P_X = \frac{1}{\sigma_X\sqrt{2\pi}} e^{-(\mu_X - x)^2/2\sigma_X{}^2}.$$

The "fuzzy" central limit theorem says that data which are influenced by many small and unrelated random effects are approximately normally distributed.