Edgeworth expansions

1 Four preliminary facts

1. You already know that $(1+a/n)^n \to e^a$. But how good is this approximation? The binomial theorem shows (after quite a bit of algebra) that for a fixed nonnegative integer k,

$$\left(1+\frac{a}{n}\right)^{n-k} = e^a\left(1-\frac{a(a+k)}{2n}\right) + o\left(\frac{1}{n}\right) \tag{1}$$

as $n \to \infty$.

2. **Hermite polynomials:** If $\phi(x)$ denotes the standard normal density function, then we define the Hermite polynomials $H_k(x)$ by the equation

$$(-1)^k \frac{d^k}{dx^k}\phi(x) = H_k(x)\phi(x). \tag{2}$$

Thus, we obtain $H_1(x) = x$, $H_2(x) = x^2 - 1$, $H_3(x) = x^3 - 3x$, and so on. By differentiating (2), we obtain

$$\frac{d}{dx}\left[H_k(x)\phi(x)\right] = -H_{k+1}(x)\phi(x). \tag{3}$$

3. **An inversion formula for characteristic functions:** Suppose $X \sim G(x)$ and $\psi_X(t)$ denotes the characteristic function of X. If $\int_{-\infty}^{\infty}|\psi_X(t)|\,dt < \infty$, then $g(x) = G'(x)$ exists and

$$g(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\psi_X(t)\,dt. \tag{4}$$

We won't prove this fact here, but its proof can be found in most books on theoretical probability.

4. **An identity involving $\phi(x)$:** For any positive integer k,

$$\begin{aligned}\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}e^{-t^2/2}(it)^k\,dt &= \frac{(-1)^k}{2\pi}\frac{d^k}{dx^k}\int_{-\infty}^{\infty} e^{-itx}e^{-t^2/2}\,dt \\ &= (-1)^k\frac{d^k}{dx^k}\phi(x) \qquad (5)\\ &= H_k(x)\phi(x), \qquad (6)\end{aligned}$$

where (5) follows from (4) since $e^{-t^2/2}$ is the characteristic function for a standard normal distribution, and (6) follows from (2).

2 The setup

Let $X_1, \ldots, X_n$ be a simple random sample from $F(x)$. Without loss of generality, suppose that $\mathrm{E}\ X_1 = 0$ and $\mathrm{Var}\ X_1 = 1$; for otherwise, we can replace each X_j by $(X_j - \mathrm{E}\ X_1)/\sqrt{\mathrm{Var}\ X_1}$ without changing anything that follows. Let

$$\gamma = \mathrm{E}\ X_j^3 \quad \text{and} \quad \tau = \mathrm{E}\ X_j^4$$

and suppose that $\tau < \infty$. We wish to study the distribution of the standardized sum

$$S_n = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} X_j.$$

The central limit theorem tells us that for every x, $P(S_n \le x) \to \Phi(x)$, where $\Phi(x)$ denotes the standard normal distribution function. But we would like a better approximation to $P(S_n \le x)$ than $\Phi(x)$, and we begin by constructing the characteristic function of S_n:

$$\psi_{S_n}(t) = \mathrm{E}\ \exp\{(it/\sqrt{n}) \sum_j X_j\} = \left[\psi_X(t/\sqrt{n})\right]^n.$$

We next use a Taylor expansion of $\exp\{itX/\sqrt{n}\}$: As $n \to \infty$,

$$\begin{aligned}
\psi_X\left(\frac{t}{\sqrt{n}}\right) &= \mathrm{E}\left\{1 + \frac{itX}{\sqrt{n}} + \frac{(it)^2X^2}{2n} + \frac{(it)^3X^3}{6n\sqrt{n}} + \frac{(it)^4X^4}{24n^2}\right\} + o\left(\frac{1}{n^2}\right) \\
&= \left(1 - \frac{t^2}{2n}\right) + \frac{(it)^3\gamma}{6n\sqrt{n}} + \frac{(it^4)\tau}{24n^2} + o\left(\frac{1}{n^2}\right).
\end{aligned}$$

If we raise this tetranomial to the nth power, most terms are $o(1/n)$:

$$\begin{aligned}
\left[\psi_X\left(\frac{t}{\sqrt{n}}\right)\right]^n &= \left[\left(1 - \frac{t^2}{2n}\right)^n + \left(1 - \frac{t^2}{2n}\right)^{n-1}\left(\frac{(it)^3\gamma}{6\sqrt{n}} + \frac{(it)^4\tau}{24n}\right)\right. \\
&\quad \left. + \left(1 - \frac{t^2}{2n}\right)^{n-2}\frac{(n-1)(it)^6\gamma^2}{72n^2}\right] + o\left(\frac{1}{n}\right). \qquad (7)
\end{aligned}$$

By equations (1) and (7) we conclude that

$$\begin{aligned}
\psi_{S_n}(t) &= e^{-t^2/2}\left[1 - \frac{t^4}{8n} + \frac{(it)^3\gamma}{6\sqrt{n}} + \frac{(it)^4\tau}{24n} + \frac{(it)^6\gamma^2}{72n}\right] + o\left(\frac{1}{n}\right) \\
&= e^{-t^2/2}\left[1 + \frac{(it)^3\gamma}{6\sqrt{n}} + \frac{(it)^4(\tau-3)}{24n} + \frac{(it)^6\gamma^2}{72n}\right] + o\left(\frac{1}{n}\right). \qquad (8)
\end{aligned}$$

3 The conclusion

Putting (8) together with (4), we obtain the following density function as an approximation to the distribution of S_n:

$$\begin{aligned} g(x) &= \frac{1}{2\pi}\left(\int_{-\infty}^{\infty} e^{-itx}e^{-t^2/2}\,dt + \frac{\gamma}{6\sqrt{n}}\int_{-\infty}^{\infty} e^{-itx}e^{-t^2/2}(it)^3\,dt\right. \\ &\quad \left.+\frac{\tau-3}{24n}\int_{-\infty}^{\infty} e^{-itx}e^{-t^2/2}(it)^4\,dt + \frac{\gamma^2}{72n}\int_{-\infty}^{\infty} e^{-itx}e^{-t^2/2}(it)^6\,dt\right). \end{aligned} \tag{9}$$

Next, combine (9) with (6) to yield

$$g(x) = \phi(x)\left(1 + \frac{\gamma H_3(x)}{6\sqrt{n}} + \frac{(\tau-3)H_4(x)}{24n} + \frac{\gamma^2 H_6(x)}{72n}\right). \tag{10}$$

By (3), the antiderivative of $g(x)$ equals

$$\begin{aligned} G(x) &= \Phi(x) - \phi(x)\left(\frac{\gamma H_2(x)}{6\sqrt{n}} + \frac{(\tau-3)H_3(x)}{24n} + \frac{\gamma^2 H_5(x)}{72n}\right) \\ &= \Phi(x) - \phi(x)\left(\frac{\gamma(x^2-1)}{6\sqrt{n}} + \frac{(\tau-3)(x^3-3x)}{24n} + \frac{\gamma^2(x^5-10x^3+15x)}{72n}\right). \end{aligned}$$

The expression above is called the second-order Edgeworth expansion. By carrying out the expansion in (8) to more terms, we may obtain higher-order Edgeworth expansions. The first-order Edgeworth expansion is

$$G(x) = \Phi(x) - \phi(x)\left(\frac{\gamma(x^2-1)}{6\sqrt{n}}\right).$$

Thus, for a symmetric distribution $F(x)$, $\gamma = 0$ and the usual (zero-order) central limit theorem approximation $\Phi(x)$ is already first-order accurate.

Incidentally, the second-order Edgeworth expansion explains why the definition of kurtosis of a distribution with mean 0 and variance 1 is $\tau - 3$.