As one example of the central limit theorem, consider the familiar Bernoulli distribution for \( n \) successes in \( m + n \) trials:

\[
p(n, m \mid q) = \frac{(n + m)!}{n! m!} q^n (1 - q)^m
\]

in the limit of large \( m, n \). Recall Stirling’s formula (see next page) that for large \( N \), \( N! \sim e^{-N} N^N \sqrt{2\pi N} \), so we can write

\[
\frac{(n + m)!}{n! m!} q^n (1 - q)^m \sim \frac{\sqrt{n + m}}{\sqrt{2\pi nm}} \frac{(n + m)^{n+m}}{n^m m^m} q^n (1 - q)^m = \frac{1}{\sigma \sqrt{2\pi}} \left( \frac{q}{n/(m + n)} \right)^n \left( \frac{1 - q}{m/(m + n)} \right)^m.
\]

To simplify this, we set \( p = n/(n + m) \), so that \( 1 - p = m/(n + m) \), and let \( \sigma = \sqrt{(n + m)p(1 - p)} \), giving

\[
p(n, m \mid q) \sim \frac{1}{\sigma \sqrt{2\pi}} \left( \frac{q}{p} \right)^n \left( \frac{1 - q}{1 - p} \right)^m = \frac{1}{\sigma \sqrt{2\pi}} e^{(n + m)(p \ln \frac{q}{p} + (1 - p) \ln \frac{1 - q}{1 - p})}.
\]

The function on the right is strongly peaked near its maximum at \( p = p_0 \), where the exponent \( (n + m)f(p) \) can be approximated in terms of the Taylor series \( f(p) \approx f(p_0) + f'(p)|_{p=p_0}(p - p_0) + \frac{1}{2} f''(p)|_{p=p_0}(p - p_0)^2 \). The point \( p_0 \) where the maximum occurs is determined by the vanishing of the first derivative \( f'(p)|_{p=p_0} = 0 \). This derivative satisfies

\[
f'(p) = \frac{d}{dp} f(p) = \frac{d}{dp} \left( p \ln \frac{q}{p} + (1 - p) \ln \frac{1 - q}{1 - p} \right) = \ln \frac{q}{p} - \ln \frac{1 - q}{1 - p},
\]

and thus vanishes at \( p = q \). We see that the value at the maximum as well vanishes, \( f(q) = 0 \), so the function is approximated by \( f(p) \approx \frac{1}{2} f''(p)|_{p=q}(p - q)^2 \) near the maximum. Taking the second derivative gives

\[
f''(p) = \frac{d}{dp} f'(p) = \frac{d}{dp} \left( \ln \frac{q}{p} - \ln \frac{1 - q}{1 - p} \right) = -\frac{1}{p} - \frac{1}{1 - p} = -\frac{1}{p(1 - p)}.
\]

Inserting \( f(p) \approx -\frac{(p-q)^2}{2q(1-q)} \) into the exponent of (1) and changing variables from \( p \) back to \( n \) shows that the Bernoulli form tends to a normal distribution in this large \( n, m \) limit:

\[
p(n, m \mid q) \sim \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{n+m}{2q(1-q)}(p - q)^2} = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(n-\mu)^2}{2\sigma^2}},
\]

where as expected \( \mu = (n + m)q \) and \( \sigma = \sqrt{(n + m)q(1 - q)} \).
To prove Stirling’s formula, \( N! \sim \sqrt{2\pi N} N^N e^{-N} \) for large \( N \), we first recall the method of steepest descent. An integral of the form,

\[
I = \int_a^b e^{Nf(x)} \, dx,
\]

when \( N \) is large, is dominated by the “critical points” \( x_0 \) of \( f \), at which \( f'(x_0) = 0 \). At any such point, we can approximate in the near vicinity of \( x_0 \),

\[
f(x) \approx f(x_0) - \frac{1}{2} |f''(x_0)| (x - x_0)^2.
\]

Assuming only a single critical point, we can then approximate

\[
I \approx e^{Nf(x_0)} \int_{-\infty}^{\infty} e^{-\frac{1}{2} N |f''(x_0)| x^2} \, dx = e^{Nf(x_0)} \sqrt{\frac{2\pi}{N|f''(x_0)|}}
\]

(1)

(\text{where we have also extended the limits of integration from } -\infty \text{ to } \infty, \text{ assuming that only the region near } x_0 \text{ matters anyway, and then used } \int_{-\infty}^{\infty} dx \, e^{-ax^2} = \sqrt{\pi/a} \, \dagger).}

To apply to the factorial function, we first need an integral representation. The function

\[
\Gamma(N + 1) = \int_0^\infty e^{-x} x^N \, dx
\]

satisfies the recursion relation

\[
\Gamma(N + 1) = \left. -e^{-x} x^N \right|_{0}^{\infty} + N \int_0^\infty e^{-x} x^{N-1} \, dx = N \Gamma(N)
\]

(using integration by parts). Together with the boundary condition \( \Gamma(1) = \int_0^\infty e^{-x} \, dx = 1 \), we see that \( \Gamma(N + 1) = N! \) for integer \( N \).

So we write

\[
N! = \Gamma(N + 1) = N^{N+1} \int_0^\infty e^{N(\ln z - z)} \, dz
\]

where \( z = x/N \). Hence \( f = \ln z - z, \, f' = 1/z - 1, \, f'' = -1/z^2 \), and \( z_0 = 1 \). Eqn. (1) gives

\[
N! \approx N^{N+1} e^{-N} \sqrt{\frac{2\pi}{N}} = \sqrt{2\pi N} N^N e^{-N}.
\]

\[\dagger\] If we write \( J = \int_{-\infty}^{\infty} dx \, e^{-x^2} \), then using \( r, \theta \) cylindrical coordinates its square can be written \( J^2 = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \, e^{-x^2-y^2} = \int_0^{2\pi} \, d\theta \int_0^{\infty} \, dr \, r \, e^{-r^2} = 2\pi \left(-\frac{1}{2} e^{-r^2} \right|_{0}^{\infty} = \pi \). Thus \( J = \sqrt{\pi} \) and \( \int_{-\infty}^{\infty} dx \, e^{-ax^2} = \sqrt{\pi/a} \).