1 Stirling's formula

Start with the definition of the continuous factorial:

$$
\begin{aligned}
n! &= \int_0^\infty x^n \cdot e^{-x} \cdot dx && (1)\\
&= \int_0^\infty e^{-x+n\log x} \cdot dx && (2)\\
&= n \cdot \int_0^\infty e^{-ny+n\log(ny)} \cdot dy && (3)
\end{aligned}
$$

With $f(y) = -y + \log(ny)$ this can be written as

$$= n \cdot \int_0^\infty e^{nf(y)} \cdot dy. \tag{4}$$

The maximum of $f(y)$ is at $y_0 = 1$. Taylor expansion around y_0:

$$f(y) \approx \log(n) - 1 - \frac{(y-1)^2}{2}. \tag{5}$$

Thus

$$
\begin{aligned}
n! &\approx ne^{n\log(n)-n} \cdot \int_0^\infty e^{-n(y-1)^2/2} \cdot dy\\
&= ne^{n\log(n)-n}\sqrt{\frac{\pi}{2n}}\left(1+\mathrm{Erf}\left(\sqrt{\frac{n}{2}}\right)\right)\\
&\approx \sqrt{2\pi n}e^{n\log(n)-n}, && (6)
\end{aligned}
$$

where we have used $\mathrm{Erf}(x) \approx 1$ for large x. Taking the logarithm yields the Stirling formula:

$$\log(n!) = n\log(n) - n + \frac{1}{2}\log(2\pi n). \tag{7}$$

2 Adding two Gaussian distributions

a) Without characteristic function

Let p_x and p_y be Gaussian distributions with parameters μ_x, μ_y and σ.

$$p_x(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(- \frac{(x-\mu_x)^2}{2\sigma^2} \right), \tag{8}$$

$$p_y(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(- \frac{(y-\mu_y)^2}{2\sigma^2} \right). \tag{9}$$

$$\begin{aligned}
p_z(z) &= \int dx \int dy \cdot \delta(z-(x+y)) \cdot p(x,y) && (10)\\
&= \int dx \int dy \cdot \delta(z-(x+y)) \cdot p(x) \cdot p(y) && (11)\\
&= \frac{1}{2\pi\sigma^2} \int dx \int dy \cdot \delta(z-(x+y)) \cdot \exp\left(- \frac{(y-\mu_y)^2}{2\sigma^2} \right) \cdot \exp\left(- \frac{(x-\mu_x)^2}{2\sigma^2} \right) && (12)\\
&= \frac{1}{\sqrt{4\pi\sigma^2}} \cdot \exp\left(- \frac{(z-\mu_x-\mu_y)^2}{4\sigma^2} \right). && (13)
\end{aligned}$$

With $\mu_z = \mu_x + \mu_y$ and $\sigma_z = \sqrt{2}\sigma$, this can be rewritten as

$$p_z(z) = \frac{1}{\sqrt{2\pi\sigma_z^2}} \cdot \exp\left(- \frac{(z-\mu_z)^2}{2\sigma_z^2} \right). \tag{14}$$

b) With characteristic function

The Fourier transform of a normal density with mean μ and standard deviation σ is

$$\hat{p}(t) = \int_{-\infty}^{\infty} p(x) e^{-itx} dx = e^{-i\mu t} e^{-\frac{(\sigma t)^2}{2}}. \tag{15}$$

The product of two distributions with the same σ is

$$\hat{p}_z(t) = \hat{p}_x(t)\hat{p}_y(t) = e^{-i(\mu_x+\mu_y)t} e^{-(\sigma t)^2}, \tag{16}$$

which leads also to the definitions $\mu_z = \mu_x + \mu_y$ and $\sigma_z = \sqrt{2}\sigma$ and a Gaussian distribution.

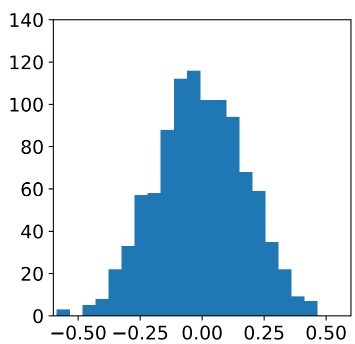
3 Computer exercise on random numbers

Write a little computer program that sums up N random numbers drawn from $[-1, 1]$. Divide by N and compare the obtained distribution to the Gaussian distribution, for $N = 10, 100, 1000$, by plotting the histogram and the appropriate analytical function together.

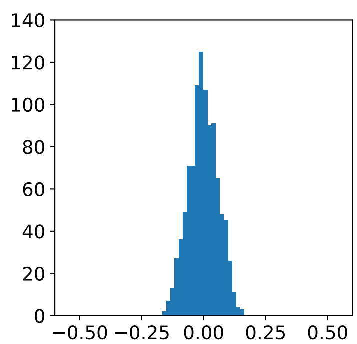
To create a distribution we run a loop for 1000 iterations, at each iteration we create a random number by summing up N number between -1 and 1. The computer program can be implemented using the following python snippet:

```python
import random

import matplotlib.pyplot as plt
import numpy as np
from scipy.optimize import curve_fit

M = 1000
for N in [10, 100, 1000]:
    rs = []
    for _ in range(M):
        r = 0
        for i in range(N):
            r += random.uniform(-1, 1)
        rs.append(r / N)

    plt.figure(figsize=(3, 3))
    plt.xlim(-0.6, 0.6)
    plt.ylim(0, 130)
    plt.hist(rs, bins=20)
    plt.savefig(f'../figures/{N}.pdf')
```

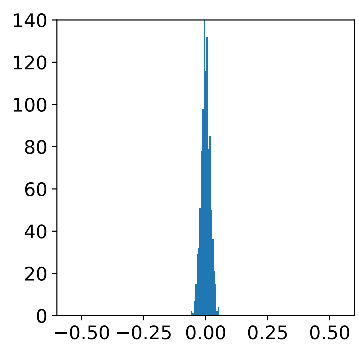
This leads to these plots:



(a) $N = 10$



(b) $N = 100$



(c) $N = 1000$

As is to be expected, the expectation value of the distrition does not change with increasing N, but stays constant at ≈ 0. The width of the curve decreases for a large number of summands. This is due to the fact that we divide the sum by N.