ENGINEERING TRIPOS PART II A

EIETL **MODULE EXPERIMENT 3F3**

RANDOM VARIABLES and RANDOM NUMBER GENERATION
Short Report Template

Name: **Abhishek Shenoy (ams289)**

College: **Pembroke College**

Lab Group Number:

1. **Uniform and normal random variables.**

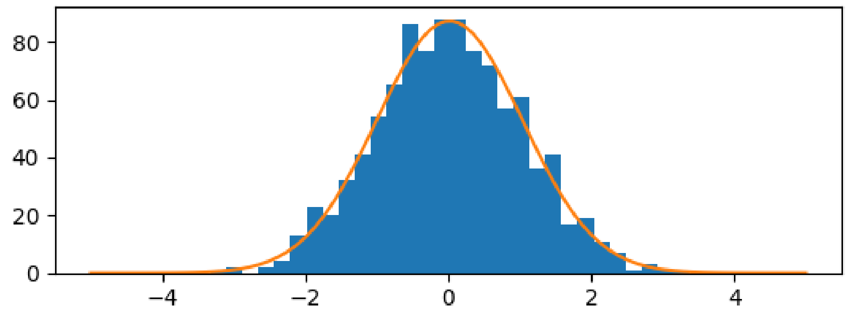


Figure 1: Histogram of Gaussian random numbers overlaid on exact Gaussian curve (scaled).

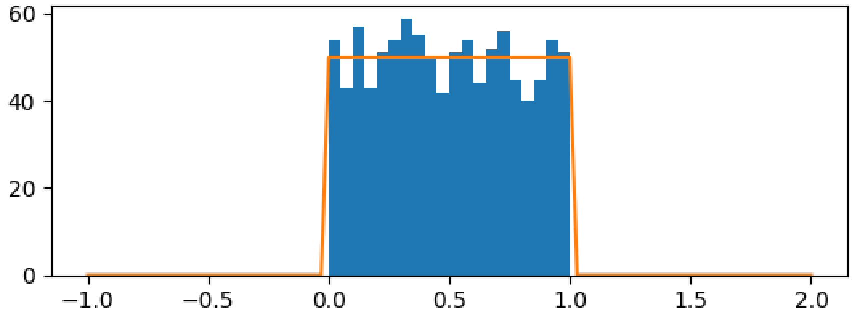


Figure 2: Histogram of Uniform random numbers overlaid on exact Uniform curve (scaled).

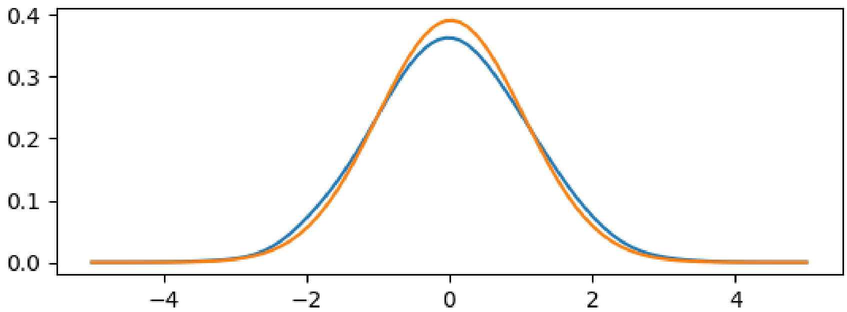


Figure 3: Kernel density estimate for Gaussian random numbers overlaid on exact Gaussian curve.

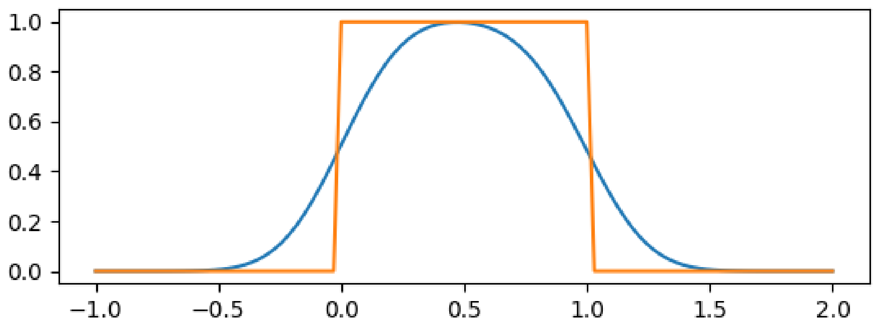


Figure 4: Kernel density estimate for Uniform random numbers overlaid on exact Uniform curve.

Comment on the advantages and disadvantages of the kernel density method compared with the histogram method for estimation of a probability density from random samples:

Kernel Density Estimation

Advantages:

- Converges quicker than a histogram for a gaussian distribution.
- More flexible as the bandwidth can be varied allowing a greater control for fitting the data.
- Working with a continuous distribution allows greater analysis and processing than discrete data.

Disadvantages:

- Converges slower than a histogram for a uniform distribution.
- KDE is very much dependent on the bandwidth and the kernel function.
- An appropriate kernel function must be chosen for a suitable estimation.

Theoretical mean and standard deviation calculation for uniform density as a function of N:

The probability that a sample x_i lies within a particular bin of the histogram is:

$$p_j = \int_{c_j-\delta/2}^{c_j+\delta/2} p(x)dx$$

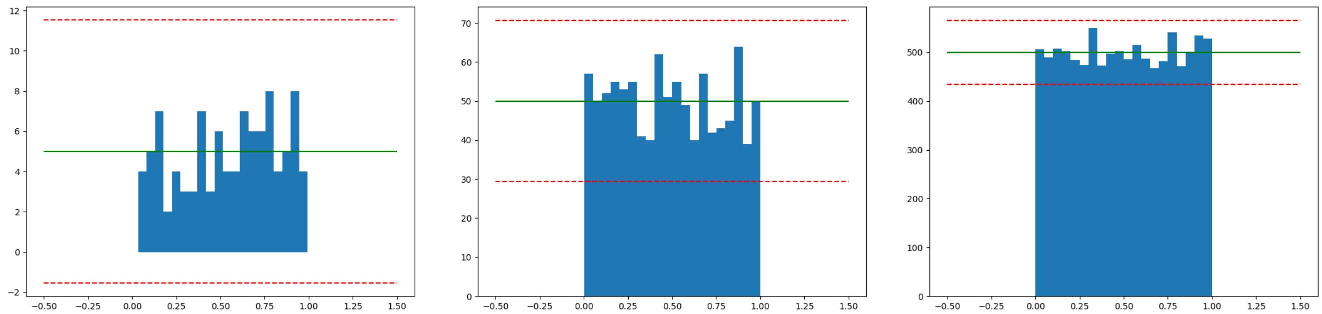
where c_j is the j-th bin centre, $j = 1, ..., J$ and δ is the bin width. If we draw N random variables and count the number that lie within each bin (n_j), the probability of the histogram data is given by the multinomial distribution:

$$\frac{N!}{n_1!n_2!...n_J!} p_1^{n_1} p_2^{n_2} p_J^{n_J}$$

The mean of the count data in bin j is Np_j and the variance is $Np_j(1 - p_j)$.
Therefore the theoretical mean for a uniform distribution is $N\delta$ and standard deviation $\sqrt{N\delta(1-\delta)}$

Explain behaviour as N becomes large:

As N becomes large, the mean increases proportionally however the standard deviations increases with $\sqrt{N}$.



(a) N = 100 (b) N = 1000 (c) N = 10000

Figure 5: Plot of histograms for $N = 100$, $N = 1000$ and $N = 10000$ with theoretical mean (green) and ± 3 standard deviation lines (red)

Are your histogram results consistent with the multinomial distribution theory?

This is consistent with the multinomial distribution theory as it is possible to see that the standard deviation grows with $\sqrt{N}$ while the mean grows with N. As N increases, the mean becomes much larger in comparison with the standard deviation. As an example, using $B = 20$ bins gives $\delta = 1/20$ and hence $\sigma = \sqrt{N}\frac{\sqrt{B-1}}{B}$ which for $N = 100$ gives $\sigma = 10\frac{\sqrt{19}}{20} \approx 2.18$ therefore $3\sigma \approx 6.54$ is actually larger than the mean 5 giving a lower bound less than zero. Similarly for $N = 1000$, $3\sigma \approx 65.4$ while the mean is 500.

2. **Functions of random variables For normally distributed $\mathcal{N}(x|0,1)$ random variables, take $y = f(x) = ax + b$. Calculate $p(y)$ using the Jacobian formula:**

For a given random variable y which is a function $f(x)$ of another random variable x with the density function $p_x(x)$, then the following is true where $f_k^{-1}(y)$ is one of K inverse functions of f:

$$p_y(y) = \sum_{k=1}^{K} \left. \frac{p_x(x)}{\left|\frac{dy}{dx}\right|} \right|_{x=f_k^{-1}(y)} \tag{1}$$

The inverse function for $y = ax + b$ is $f^{-1}(y) = \frac{y-b}{a}$ and $\frac{dy}{dx} = a$, substituting these into equation 1, we get the following:

$$\begin{aligned} p_y(y) &= \frac{p_x(\frac{y-b}{a})}{a} \\ &= \frac{1}{\sqrt{2\pi a^2}} e^{-(y-b)^2/2a^2} = \mathcal{N}(y|b,a) \end{aligned}$$

Explain how this is linked to the general normal density with non-zero mean and non-unity variance:

Therefore $p_y(y)$ is a normal distribution with mean $\mu = b$ and $\sigma = a$.

Verify this formula by transforming a large collection of random samples $x^{(i)}$ to give $y^{(i)} = f(x^{(i)})$, histogramming the resulting y samples, and overlaying a plot of your formula calculated using the Jacobian:

Letting $a = 2, b = 5$:

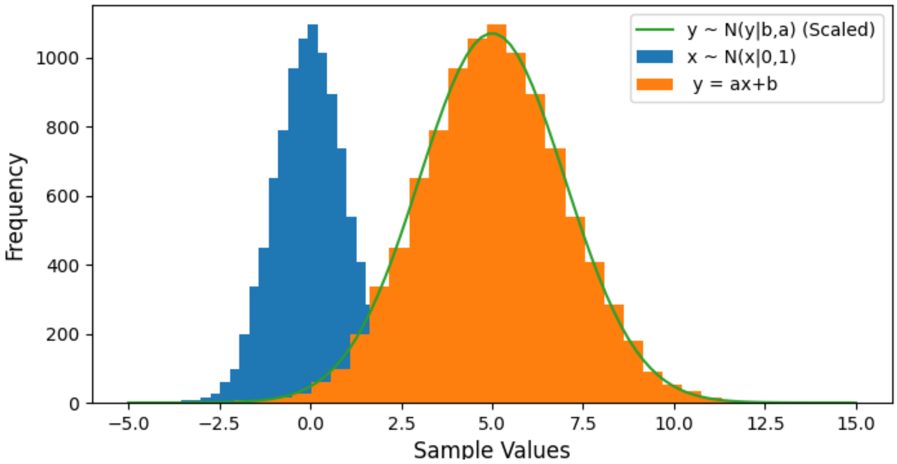


Figure 6: Plot of histograms for x and y with an analytical solution for y.

Now take $p(x) = \mathcal{N}(x|0,1)$ and $f(x) = x^2$. Calculate $p(y)$ using the Jacobian formula:

$$x = \pm\sqrt{y} \quad \frac{dy}{dx} = 2x$$

$$\begin{aligned}
p_y(y) &= \sum \frac{p_x(x)}{|\frac{dy}{dx}|} \\
&= \frac{p_x(\sqrt{y})}{|2x|} + \frac{p_x(-\sqrt{y})}{|2x|} \\
&= \frac{p_x(\sqrt{y})}{|2\sqrt{y}|} + \frac{p_x(-\sqrt{y})}{|-2\sqrt{y}|} \\
&= 2 \cdot \frac{p_x(\sqrt{y})}{|2\sqrt{y}|} \\
&= \frac{p_x(\sqrt{y})}{\sqrt{y}} \\
p_y(y) &= \frac{1}{\sqrt{2\pi y}} e^{-y/2}
\end{aligned}$$

Verify your result by histogramming of transformed random samples:

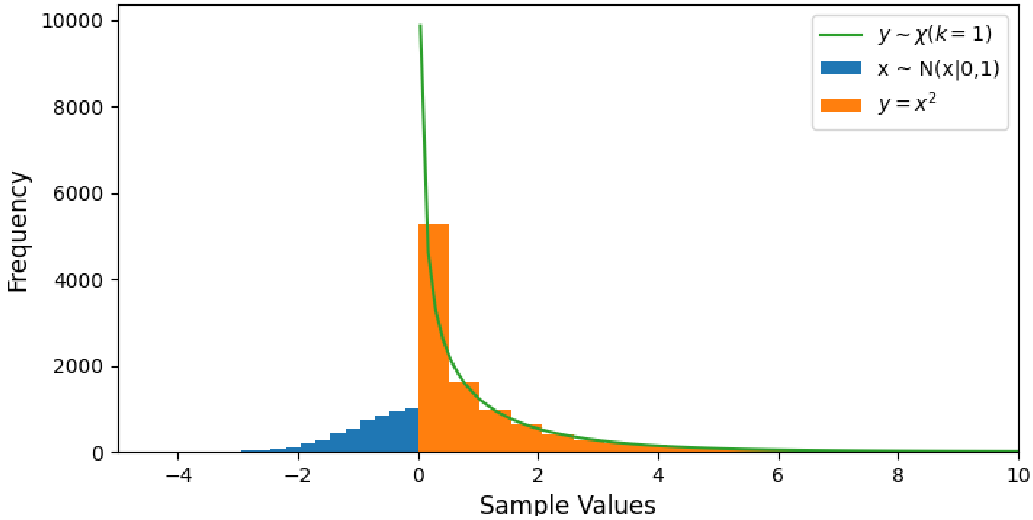


Figure 7: Plot of histograms for x and y with an analytical solution for y.

3. **Inverse CDF method**

   Calculate the CDF and the inverse CDF for the exponential distribution:

   Given the exponential distribution $p(x) = e^{-x}$:

$$\begin{aligned} F(x) &= \int_0^x p(x)dx \\ &= \int_0^x e^{-x}dx \\ &= [-e^{-x}]_0^x \\ &= 1 - e^{-x} \end{aligned}$$

$$F^{-1}(x) = -ln(1-x)$$

Plot histograms/ kernel density estimates and overlay them on the desired exponential density:

```python
import numpy as np
import matplotlib.pyplot as plt
import scipy.stats

N, numbins = 10000, 50
x = np.random.rand(N)
# Inverse CDF for exponential distribution
u = -np.log(1-x)

fig, ax = plt.subplots(1,figsize=(8,4))
# Plot histograms
ax.hist(u, bins=numbins)
ax.hist(x, bins=numbins)

# Plot analytical distribution
x_values = np.linspace(0., 8., 100)
y_values = np.exp(-x_values)
scale = N * (u.max() - u.min())/numbins
ax.plot(x_values, y_values*scale)

plt.xlim([-2,8])
plt.xlabel("Sample Values", fontsize=12)
plt.ylabel("Frequency", fontsize=12)
plt.legend(["$y = F^{-1}_{e^{-x}}(x) $", "x ~ U(0,1)", "$e^{-x}$ (scaled)"])
plt.show()

```

Figure 8: Code for inverse CDF method for generating samples from the exponential distribution

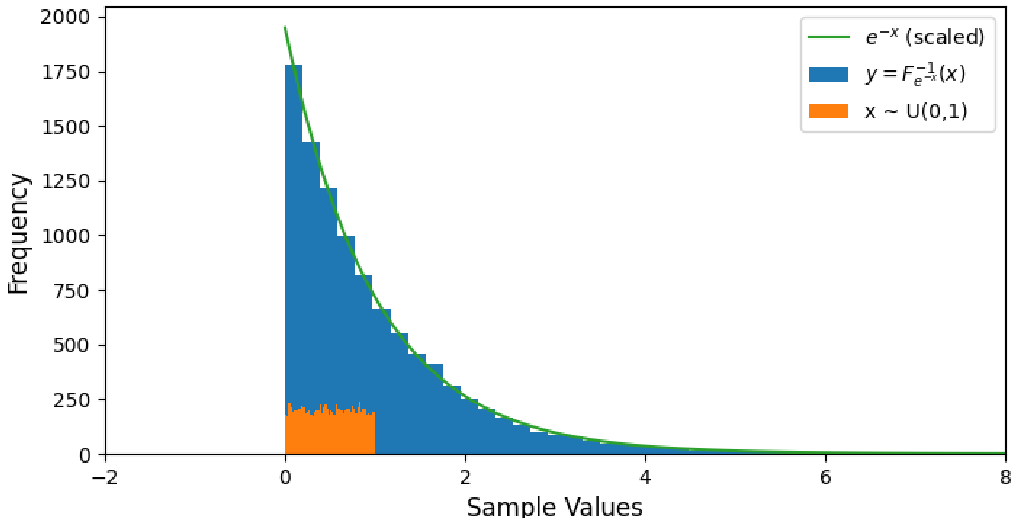


Figure 9: Histogram for the inverse CDF method on the exponential distribution with the desired analytical p.d.f

4. **Simulation from a 'non-standard' density.**

$$p(x) = \int_0^\infty \mathcal{N}(x|0, u)p(u)du$$

where

$$p(u) = \frac{\alpha^2}{2} e^{-\frac{\alpha^2}{2}u}$$

```
import numpy as np

a = 1.0
N = 10000
pdf_new = []
for i in range(N):
    z = np.random.rand()
    # Inverse CDF for exponential distribution
    u = (-2/(a**2)) * np.log(1-z)
    # u is the variance of the normal distribution
    x_given_u = np.random.normal(0, np.sqrt(u))
    x = x_given_u * u
    pdf_new.append(x)
```

Figure 10: Code to generate N random numbers drawn from the distribution of X

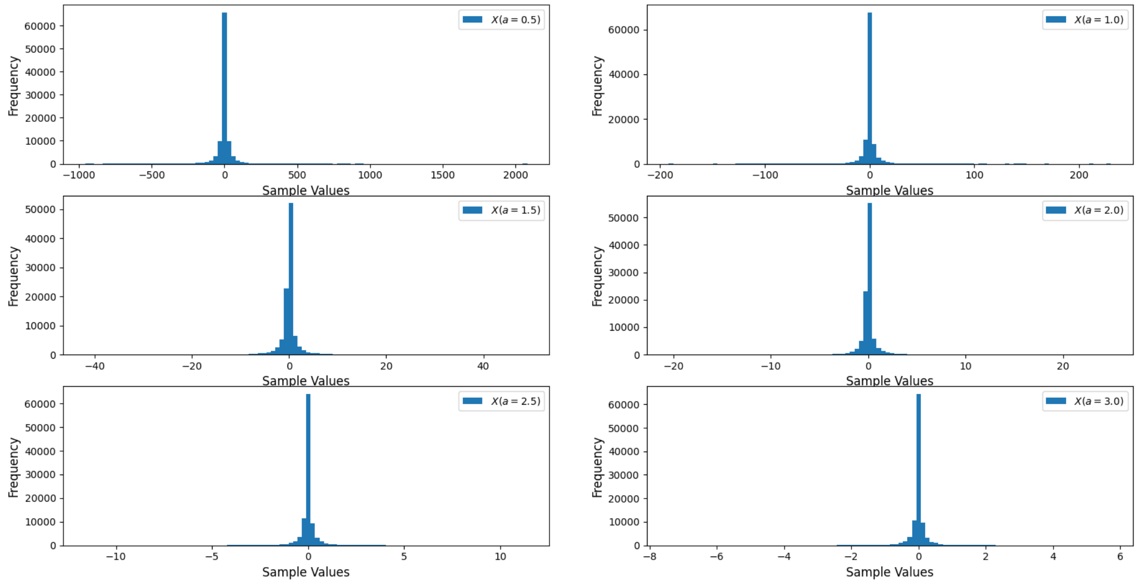


Figure 11: Histogram density estimates for the distribution of X for different values of α

Hence comment on the interpretation of the parameters α and its effect on the density of X

α decreases the spread of values exponentially. This is possible to see by looking at the values on the x-axis of Fig. 11 and noticing that the scale decreases as α increases. Therefore as α increases, the density becomes more and more concentrated at the origin.

Plot the logarithm of the kernel density estimate (be careful around the origin) for different values of α and use this to suggest a formula for p(x). Compare the shape of the distribution to that of a Normal distribution.

p(x) appears to be a Laplace distribution as given by the sharp peak arising from the exponential distribution and the symmetry arising from the normal distribution as can be seen in Fig. 13. As α decreases, the spread of the distribution increases hence reducing the probability density around the origin. Comparing to a normal distribution, it is possible to see that the normal distribution has thinner tails than the Laplace distribution which is as expected.

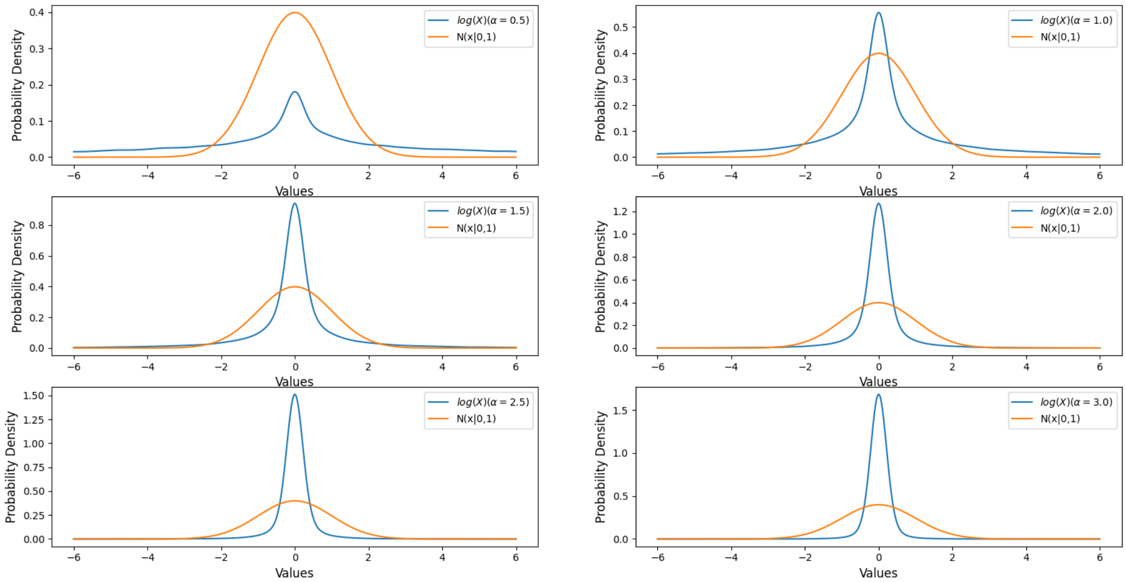


Figure 12: Kernel density estimate for the log of the distribution of X for $0.5 \leq \alpha \leq 3.0$ (corresponding to the histograms) plotted with the standard normal distribution

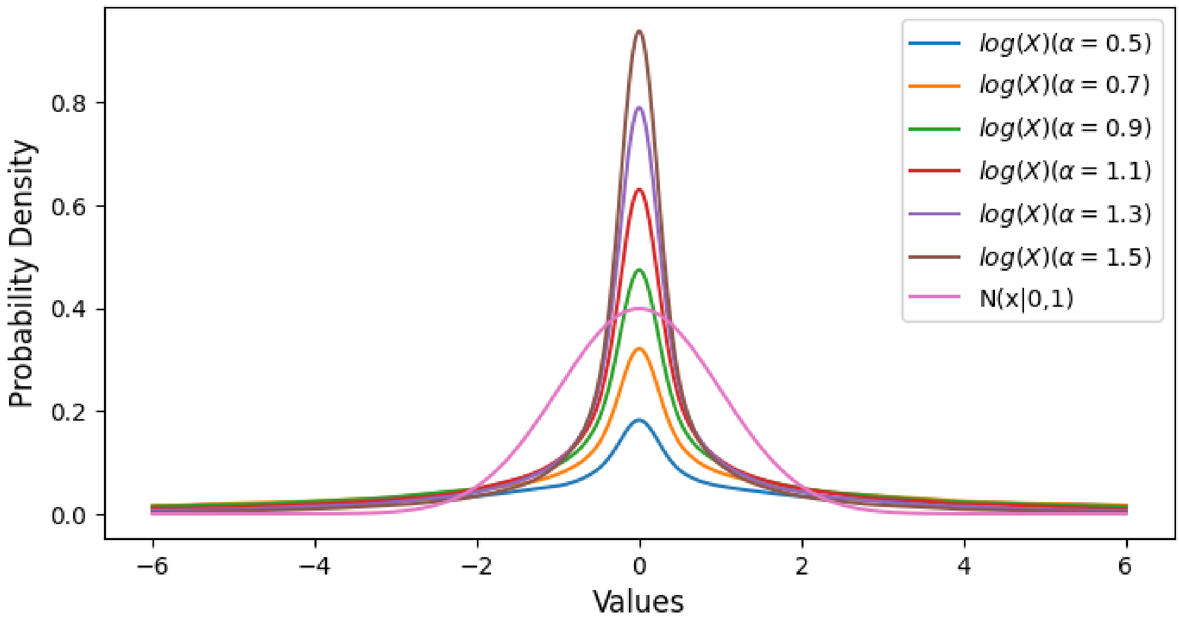


Figure 13: Overlay of all kernel density estimate graphs for $\log X$ for $0.5 \leq \alpha \leq 1.5$ plotted with the standard normal distribution

We can derive the distribution using our original statement:

$$p(x) = \int_0^\infty \mathcal{N}(x|0,u)p(u)du$$

where

$$p(u) = \frac{\alpha^2}{2}e^{-\dfrac{\alpha^2}{2}u}$$

$$\begin{aligned}
p(x) &= \int_0^\infty \mathcal{N}(x|0,u)p(u)du \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi u}}e^{-\frac{1}{2}\frac{x^2}{u}} \cdot \frac{\alpha^2}{2}e^{-\frac{\alpha^2}{2}u}du \\
&= \frac{\alpha^2}{2}\int_0^\infty \frac{1}{\sqrt{2\pi u}}e^{-\frac{\alpha^2}{2u}(u^2+\frac{x^2}{\alpha^2})}du \\
&= \frac{\alpha^2}{2}\int_0^\infty \frac{1}{\sqrt{2\pi u}}e^{-\frac{\alpha^2}{2u}\left((u-\frac{|x|}{\alpha})^2+\frac{2u}{\alpha}|x|\right)}du \\
&= \frac{\alpha^2}{2}\int_0^\infty \frac{1}{\sqrt{2\pi u}}e^{-\frac{\alpha^2}{2u}\left(u-\frac{|x|}{\alpha}\right)^2}e^{-\alpha x}du \\
&= \frac{\alpha^2}{2}e^{-\alpha x}\int_0^\infty \frac{1}{\sqrt{2\pi u}}e^{-\frac{\alpha^2}{2u}\left(u-\frac{|x|}{\alpha}\right)^2}du \\
&= \frac{\alpha^2}{2}e^{-\alpha x}\int_0^\infty \frac{u}{\sqrt{\lambda}}\sqrt{\frac{\lambda}{2\pi u^3}}e^{-\frac{\lambda}{2\mu^2}\dfrac{\left(u-\mu\right)^2}{u}}du \quad \left(\text{Rewrite as inverse gaussian pdf where } \frac{\lambda}{2\mu^2} = \frac{\alpha^2}{2} \text{ and } \mu = \frac{|x|}{\alpha}\right) \\
&= \frac{\alpha^2}{2}e^{-\alpha x}\frac{1}{\sqrt{\lambda}}\int_0^\infty u\sqrt{\frac{\lambda}{2\pi u^3}}e^{-\frac{\lambda}{2\mu^2}\dfrac{\left(u-\mu\right)^2}{u}}du \quad \left(\lambda = |x|^2 \text{ and } \mu = \frac{|x|}{\alpha}\right) \\
&= \frac{\alpha^2}{2}e^{-\alpha x}\frac{1}{\sqrt{\lambda}}E(\text{InvGauss}(\lambda,\mu)) \\
&= \frac{\alpha^2}{2}e^{-\alpha x}\frac{1}{\sqrt{\lambda}}\mu \\
p(x) &= \frac{\alpha}{2}e^{-\alpha|x|}
\end{aligned}$$

This is indeed a form of the Laplace distribution.