

# CAMBRIDGE UNIVERSITY ENGINEERING DEPARTMENT

## Part IIA Full Technical Report

### 3F3 Random Number Generation

Name: Yongqing Jiang

CRSid: yj375

College: Peterhouse

Date of Experiment: Nov. 2025

## 1 Introduction

This lab activity investigates statistical methods by generating random numbers with underlying distribution. Uniform and normal random variables are generated, and they are visualized by histogram and kernel smoothing density (KSD) function. Functions of random variables are also discussed. Inverse CDF method is used to generate random variables from any arbitrary distributions. Finally, stable distribution is discussed.

## 2 Methods, Results, and Discussion

### 2.1 Uniform and normal random variables

In this section, uniform and normal random variables are generated and visualized by using histogram and kernel density function. The results are compared and discussed.

#### 2.1.1 Histogram and kernel density function

In Figure 1, 1000 random Gaussian numbers and 1000 random uniform numbers are generated. The histograms are generated with 50 bins and the KSD functions has a width of 0.4 with a

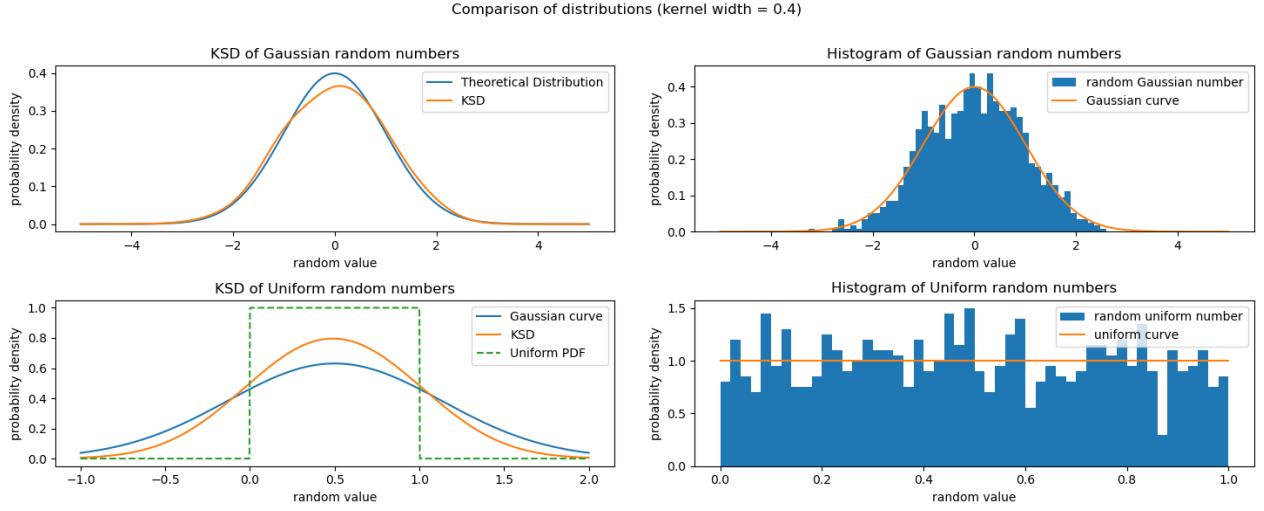


Figure 1: Histogram and KSD of uniform and normal random variables

Gaussian kernel, in the form of[1]:

$$\pi_{KS}(x) = \frac{1}{N} \sum_{i=1}^N \frac{1}{\sigma} \mathcal{K}\left(\frac{x - x^{(i)}}{\sigma}\right)$$

Where  $\mathcal{K}(\cdot) \sim \mathcal{N}(\cdot | 0, 1)$ ,  $\sigma$  is the width of the kernel.

For Gaussian random variables, the KSD provides a smooth approximation, which closely follows the shape of the theoretical Gaussian distribution curve. The histogram, on the other hand, shows more details and fluctuations due to the discrete nature of the bins.

At the same time, the histogram may not align with the theoretical curve perfectly due to the choice of bin width and the number of samples.

On the other hand, the uniform random variables show a different behavior in histogram and KSD. The shape of the histogram is more likely to be affected by the bin width, leading to a step-like appearance. Also, it is prone to be affected by random fluctuations.

The KSD, however, smooths out these fluctuations and provides a more continuous representation of the uniform distribution. At the same time, it can lead to deviation from the ideal uniform shape, due to the discontinuities at the edges of the uniform distribution. This is because the kernel width is comparable to the range of the uniform distribution.

To discuss more about the effect of kernel width, several KSDs with distinct kernel widths are plotted.

From Figure 2, we can see that with a small kernel width (0.05 and 0.1), the KSD captures more details of the uniform distribution. It is reasonable to state that this is a uniform distribution only by looking at the KSD. However, with a larger kernel width that is comparable to the range of the uniform distribution , the KSD smooths out the details and deviates from the ideal uniform shape. To conclude, there is a trade-off when choosing the kernel width between smoothing and shape deviation.

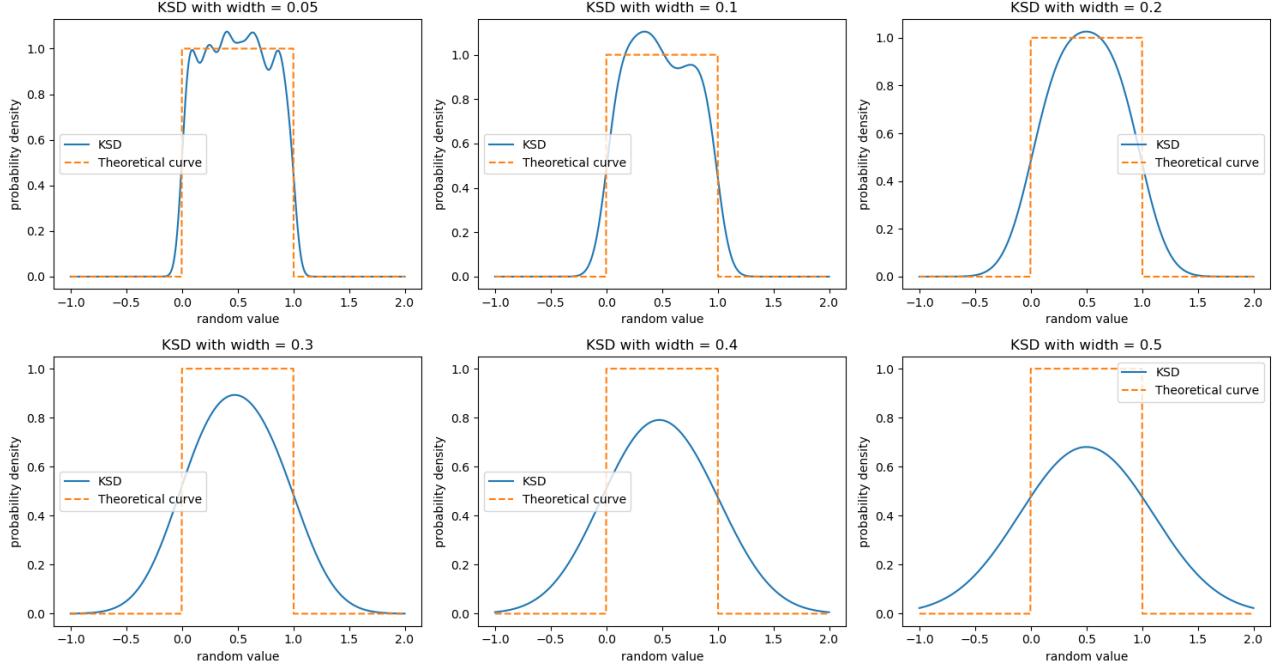


Figure 2: KSD of uniform random variables with different kernel widths

### 2.1.2 Multinomial distribution

The Multinomial distribution is a generalization of the binomial distribution [2], and it can be used to describe a distribution within a histogram.

Suppose there are some fixed finite number of bins,  $J$ , and there are  $N$  samples to be placed into these bins. Each bin has a probability  $p_j$  of receiving a sample, where  $j = 1, 2, \dots, J$  and  $\sum_{j=1}^J p_j = 1$ .

And  $p_j$  can be calculated by integrating the underlying probability density function  $p(x)$ .

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

Where  $c_j$  is the center of bin  $j$ , and  $\delta$  is the width of each bin.

Let random variable  $\vec{X} = (X_1, X_2, \dots, X_J)$  represents the number of samples in each bin. Then the PMF of the Multinomial distribution is given by:

$$P(\vec{X} = \vec{n}) = \frac{N!}{n_1! n_2! \dots n_J!} p_1^{n_1} p_2^{n_2} \dots p_J^{n_J}$$

Where  $\vec{n} = (n_1, n_2, \dots, n_J)$ . When  $\dim \vec{X}$  is 2, it reduced to a binomial distribution.

The term  $p_j^{n_j}$  represents the probability of  $n_j$  samples falling into bin  $j$ , as each sample has a probability  $p_j$  of falling into that bin.

The factorial terms account for the different arrangements of samples across the bins, considering that the order of samples within each bin does not matter.

Expectation of  $X_j$  is given by:

$$\mathbb{E}[X_j] = Np_j$$

And Variance of  $X_j$  is given by:

$$\text{Var}(X_j) = Np_j(1 - p_j)$$

For uniform distributed random variables,  $p(x) = \frac{1}{N}$ , so  $p_j = \frac{\delta}{N}$ .  
 $\mathbb{E}(X_j) = \frac{N\delta}{N} = \delta$ ,  $\text{Var}(X_j) = N\frac{\delta}{N}(1 - \frac{\delta}{N}) = \delta(1 - \frac{\delta}{N})$ .

Three sets of uniform random variables are generated with  $N = 100, 1000, 10000$  respectively. The histograms are plotted with 30 bins, and the results are shown in Figure 3.

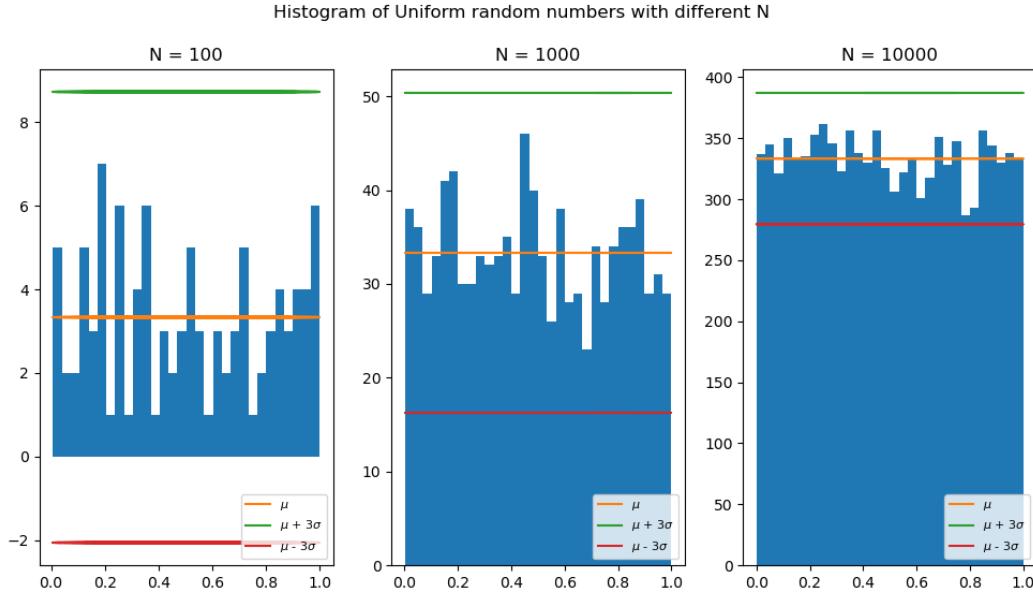


Figure 3: Histograms of uniform random variables with different sample sizes

We can see that the normalized mean of each bin (height of each bar) is fixed by the bin width, which is  $\frac{1}{30} \approx 0.0333$ .

For  $N = 1000$  and  $10000$ , the histogram bars fluctuate around this value and fits in the interval  $[\mu - 3\sigma, \mu + 3\sigma]$ , as expected.

However, for  $N = 100$ , the histogram bars show significant deviation from the expected range. This is because with a small sample size, the variance is relatively large compared to the mean. It is implied that a small sample size may not accurately represent the underlying distribution by using a histogram.

This analysis can also be applied to Gaussian random variables.

To start with,

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

Where  $F(x)$  is the CDF of the Gaussian distribution. This expression helps to evaluate  $p_j$  in *python* using `scipy.stats.norm.cdf` function.

Then,  $\mathbb{E}(X_j)$  and  $\text{Var}(X_j)$  can be calculated accordingly.

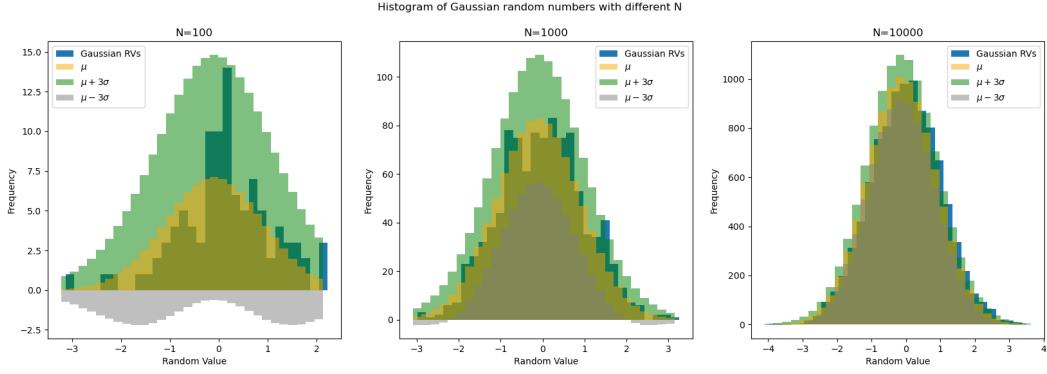


Figure 4: Histograms of Gaussian random variables with different sample sizes

As shown in Figure 4, the histograms of Gaussian random variables are plotted with corresponding expected mean and variance for each bin height. Similar to the uniform distribution case, it is expected that the bar heights lie within the interval  $[\mu - 3\sigma, \mu + 3\sigma]$ .

However, in this case, it can be observed that  $\sigma$  is no longer a constant throughout the range of random variables, since  $p_j$  varies for different bins.  $\text{Var}(X_j) = Np_j(1 - p_j)$ . The variance is larger at the center of the distribution where  $p_j$  is larger, and smaller at the tails where  $p_j$  approaches to 0 and 1.

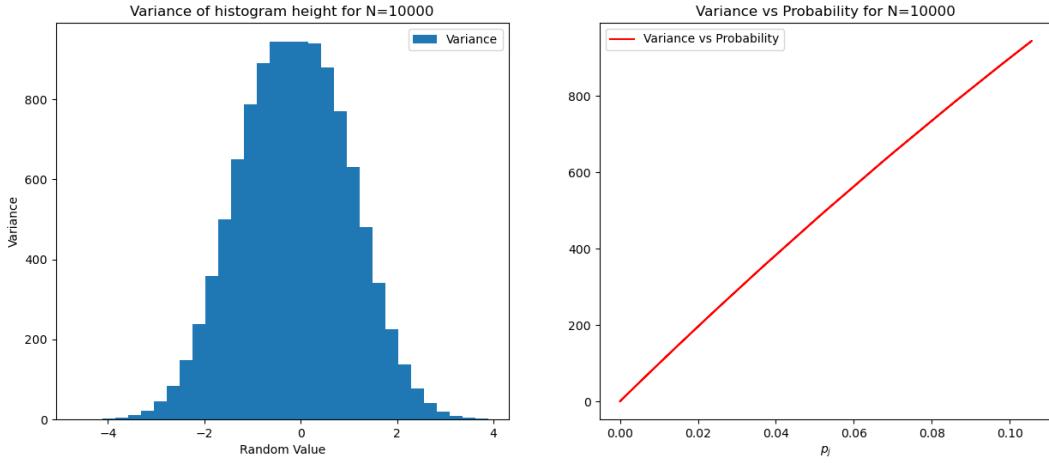


Figure 5: Comparing variance of histogram height and  $p_j$

This argument is visualized in Figure 5, where the variance of histogram height and  $p_j$  are plotted together. It is clear that the variance is small when  $p_j$  is small and at the tails.

Due to the limited bin size,  $\max(p_j)$  is around 0.1. In this range,  $\text{Var}(X_j) \approx Np_j$ , which agrees with the linear trend shown.

### 3 Function of Random Variables

The Jacobian formula for change of variables in probability density functions states that if  $y = f(x)$  is a differentiable and invertible function, then the probability density function of the transformed random variable  $y$  is given by:

$$p(y) = \frac{p(x)}{|dy/dx|} \Big|_{x=f^{-1}(y)}$$

For a linear transformation,  $y = f(x) = ax + b \Rightarrow x = f^{-1}(y) = \frac{y-b}{a}$ , and  $dy/dx = a$ . With  $f(\cdot)$  being the standard Gaussian distribution:

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$p(y) = \frac{1}{|a|} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-b)^2}{2a^2}} = \frac{1}{\sqrt{2\pi a^2}} e^{-\frac{(y-b)^2}{2a^2}}$$

For a general normal distribution  $\mathcal{N}(x|\mu, \sigma^2)$ , the probability density function is:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

We can see that the transformed random variable is  $\mathcal{N}(y|b, a^2\sigma^2)$ , the linear transformation of a normal distribution is still a normal distribution, with a shift of mean from  $\mu$  to  $\mu + b$  and a scaling of variance from  $\sigma^2$  to  $a^2\sigma^2$ .

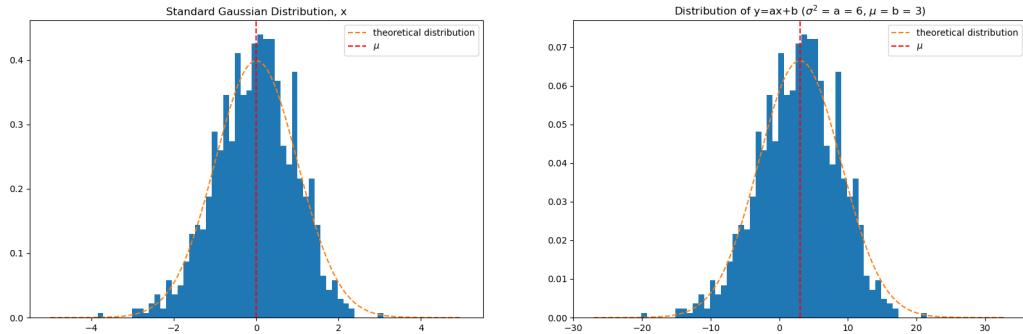


Figure 6: Histogram of linearly transformed Gaussian random variables and corresponding PDF

We can see from Figure 6 that the histogram of transformed random samples with  $a = 6$  and  $b = 3$  fits well with the probability density function with variance 6 and mean 3.

Now, consider a non-linear transformation,  $y = f(x) = x^2$ .

$$x = f^{-1}(y) = \begin{cases} \sqrt{y}, & x \geq 0 \\ -\sqrt{y}, & x \leq 0 \end{cases}$$

We can see that  $f(\cdot)$  is not one-to-one, so we need to consider both branches of the inverse function. Also,  $|dy/dx| = 2|x| = 2\sqrt{y}$ .

$$\begin{aligned} p(y) &= \frac{p(x)}{|dy/dx|} \Big|_{x=f^{-1}(y)} + \frac{p(x)}{|dy/dx|} \Big|_{x=-f^{-1}(y)} \\ &= \frac{1}{2\sqrt{y}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(\sqrt{y})^2}{2}} + \frac{1}{2\sqrt{y}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(-\sqrt{y})^2}{2}} \\ &= \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}} \end{aligned}$$

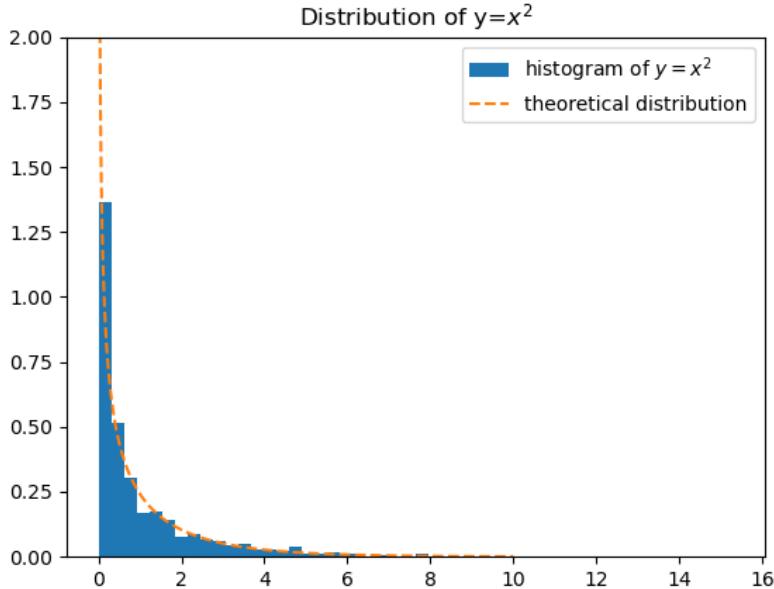


Figure 7: Histogram of squared Gaussian random variables and corresponding PDF

As shown by Figure 7, the theory agrees well to the histogram of squared Gaussian random variables.

Lastly, consider  $p(x) = \mathcal{U}(x|0, 2\pi)$ , and transformation  $y = f(x) = \sin(x)$ . The inverse of  $f(\cdot)$  is also not one-to-one within the range of  $x$ .

There are four branches of the inverse function within  $[0, 2\pi]$ , and it is better to defind Arcsin( $\cdot$ )

function that maps  $[0, 1] \rightarrow [0, \frac{\pi}{2}]$ .

$$x = f^{-1}(y) = \begin{cases} \text{Arcsin}(y), & 0 \leq x \leq \frac{\pi}{2}, 0 \leq y \leq 1 \\ \pi - \text{Arcsin}(y), & \frac{\pi}{2} \leq x \leq \pi, 0 \leq y \leq 1 \\ \pi + \text{Arcsin}(-y), & \pi \leq x \leq \frac{3\pi}{2}, -1 \leq y \leq 0 \\ 2\pi - \text{Arcsin}(-y), & \frac{3\pi}{2} \leq x \leq 2\pi, -1 \leq y \leq 0 \end{cases}$$

Also,  $|dy/dx| = |\cos(x)|$ .

Now, the transformed PDF can be calculated as:

$$p(y) = \sum_{i=1}^2 \frac{p(x)}{|dy/dx|} \Big|_{x=f_i^{-1}(y)} = \sum_{i=1}^2 \frac{1}{2\pi|\cos(x)|} \Big|_{x=f_i^{-1}(y)}, \quad 0 \leq y \leq 1$$

$|\cos(x)| = \sqrt{1 - y^2}$  for all four branches of  $f^{-1}(\cdot)$ , so

$$p(y) = \frac{2}{2\pi\sqrt{1-y^2}} = \frac{1}{\pi\sqrt{1-y^2}}, \quad 0 \leq y \leq 1$$

Similarly,  $p(y) = 1/(\pi\sqrt{1-y^2})$ , for  $-1 \leq y \leq 0$ . Combining both parts, we have:

$$p(y) = \frac{1}{\pi\sqrt{1-y^2}}, \quad -1 \leq y \leq 1$$

As shown by Figure 8, the histogram of  $y = \sin(x)$  agrees well with the derived PDF.

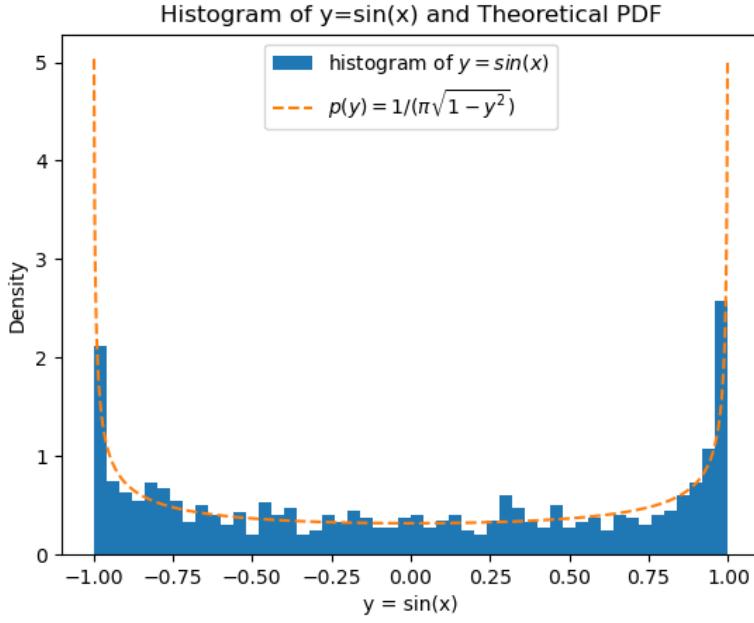


Figure 8: Histogram of  $y = \sin(x)$  and corresponding PDF

## 4 Inverse CDF Method

To generate random variables from an arbitrary distribution  $p(y)$ , the inverse CDF method is used.

Given with a uniform random number generator  $p(x) = \mathcal{U}(x|0, 1)$ . If  $y = f(x)$  is chosen to be the inverse of the CDF of  $p(y)$ ,  $F(y)$ , the generated random numbers  $y^{(i)} = f(x^{(i)})$  will follow the distribution  $p(y)$ .

This method is verified by generating exponential random variables.

The PDF of an exponential distribution Y with mean 1 is:

$$p(y) = e^{-y}$$

The corresponding cdf is found by integration:

$$F(y) = \int_0^y f_Y(t)dt = \int_0^y e^{-t}dt = 1 - e^{-y}$$

The inverse of this function is:  $F^{-1}(x) = -\ln(1-x)$ . Random numbers  $y^{(i)}$  can be generated by  $y^{(i)} = F^{-1}(x^{(i)})$ , where  $x^{(i)}$  are uniform random numbers between 0 and 1.

As shown in Figure 9, the histogram of generated exponential random variables fits well with the theoretical PDF, as well as the KSD, which verifies the correctness of the inverse CDF method for large random values.

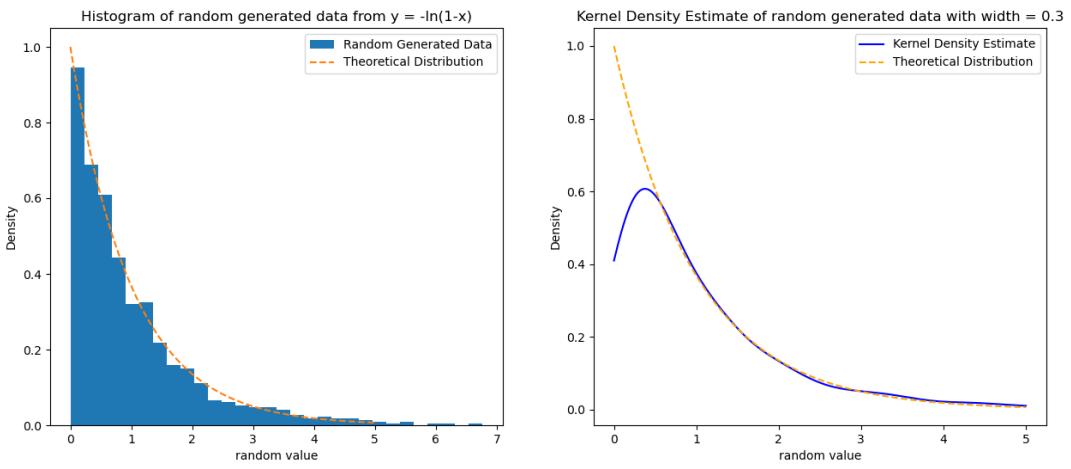


Figure 9: Histogram of exponential random variables generated by inverse CDF method and corresponding PDF

A Monte Carlo simulation is also performed to estimate the mean and variance of the exponential distribution.

$$\mu = \mathbb{E}[Y] \approx \frac{1}{N} \sum_{i=1}^N y^{(i)} = \hat{\mu}$$

$$\sigma^2 = \text{Var}(Y) \approx \frac{1}{N} \sum_{i=1}^N (y^{(i)})^2 - (\hat{\mu})^2 = \hat{\sigma}^2$$

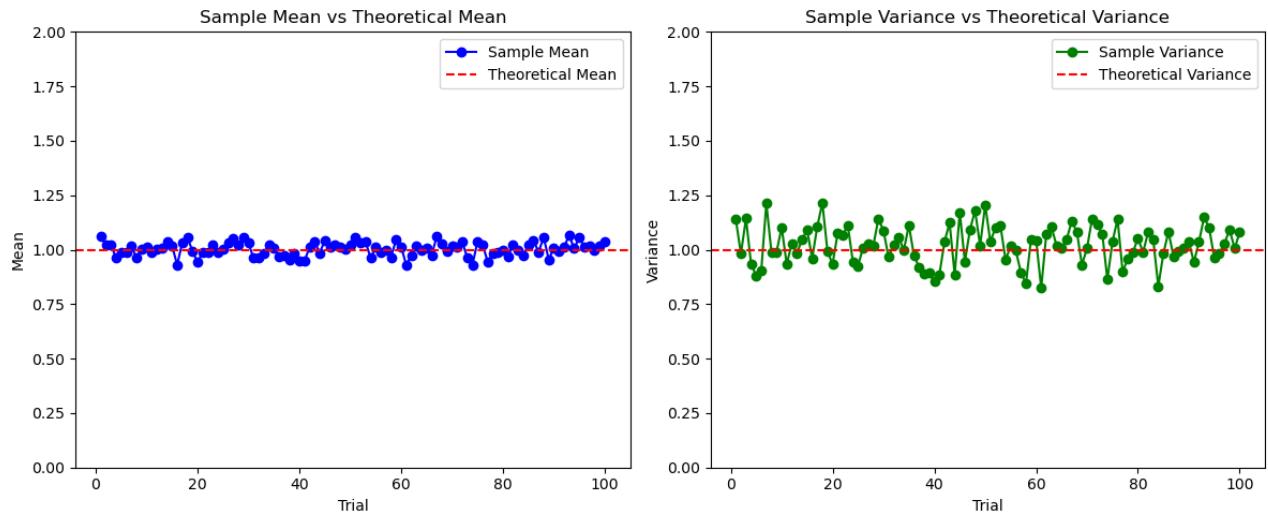


Figure 10: Mean and Variance estimation of exponential distribution by Monte Carlo simulation

100 tests are performed, and the results are shown in Figure 10. We can see that the estimators of mean and variance agrees well to the theoretical values.

The expectation of the estimator of mean  $\mathbb{E}[\hat{\mu}] = \mathbb{E}\left[\frac{1}{N} \sum_{i=1}^N y^{(i)}\right] = \frac{1}{N} \sum_{i=1}^N \mathbb{E}[y^{(i)}]$  by linearity of expectation.  $\mathbb{E}[y^{(i)}] = \mu$  by definition, so  $\mathbb{E}[\hat{\mu}] = \mu$ . Which implies that the Monte Carlo estimator of mean is unbiased.

The variance of the estimator of mean can be evaluated from the definition of variance:  $\text{Var}(\hat{\mu}) = \mathbb{E}[\hat{\mu}^2] - (\mathbb{E}[\hat{\mu}])^2 = \mathbb{E}[\hat{\mu}^2] - \mu^2$ .

$$\hat{\mu}^2 = \frac{1}{N^2} \left( \sum_{i=1}^N \sum_{j=1}^N y^{(i)} y^{(j)} \right)$$

The expecation of  $\hat{\mu}^2$  can be evaluated as:

$$\begin{aligned} \mathbb{E}[\hat{\mu}^2] &= \frac{1}{N^2} \left( \sum_{i=1}^N \sum_{j=1}^N \mathbb{E}[y^{(i)} y^{(j)}] \right) \\ &= \frac{1}{N^2} \left( \sum_{i=1}^N \mathbb{E}[y^{(i)2}] + 2 \sum_{1 \leq i < j \leq N} \mathbb{E}[y^{(i)} y^{(j)}] \right) \\ &= \frac{1}{N^2} \left( \sum_{i=1}^N \mathbb{E}[y^{(i)2}] + 2 \sum_{1 \leq i < j \leq N} \mathbb{E}[y^{(i)}] \mathbb{E}[y^{(j)}] \right) \end{aligned}$$

Linearity of expectation and independence of samples are used in the last step.

$\mathbb{E}[y^{(i)2}] = \text{Var}(Y) + (\mathbb{E}[Y])^2 = \sigma^2 + \mu^2$  and  $\mathbb{E}[y^{(i)}] \mathbb{E}[y^{(j)}] = \mu^2$  by definition. Thus,  $\mathbb{E}[\hat{\mu}^2]$  can be simplified as:

$$\begin{aligned} \mathbb{E}[\hat{\mu}^2] &= \frac{1}{N^2} \left( N(\sigma^2 + \mu^2) + 2 \frac{N(N-1)}{2} \mu^2 \right) \\ &= \frac{\sigma^2}{N} + \mu^2 \end{aligned}$$

$$\text{Var}(\hat{\mu}) = \mathbb{E}[\hat{\mu}^2] - \mu^2 = \frac{\sigma^2}{N}.$$

A final simulation is performed to verify the variance of the estimator of mean. As shown in Figure 11, the MSE of the estimator of mean decreases with increasing sample size N. For each N, 100 tests are performed to evaluate the MSE by Monte Carlo simulation.

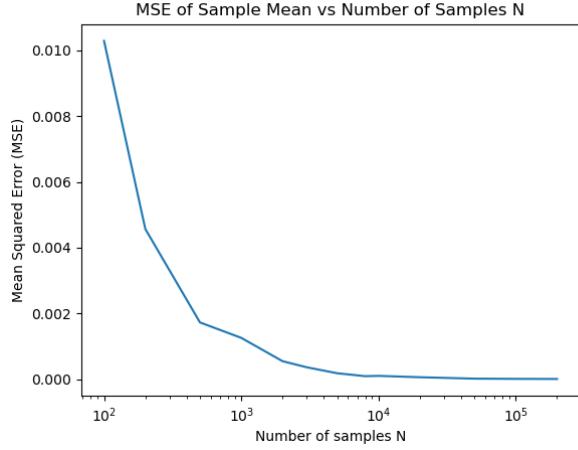


Figure 11: MSE against sample size N

## 5 Stable Distribution

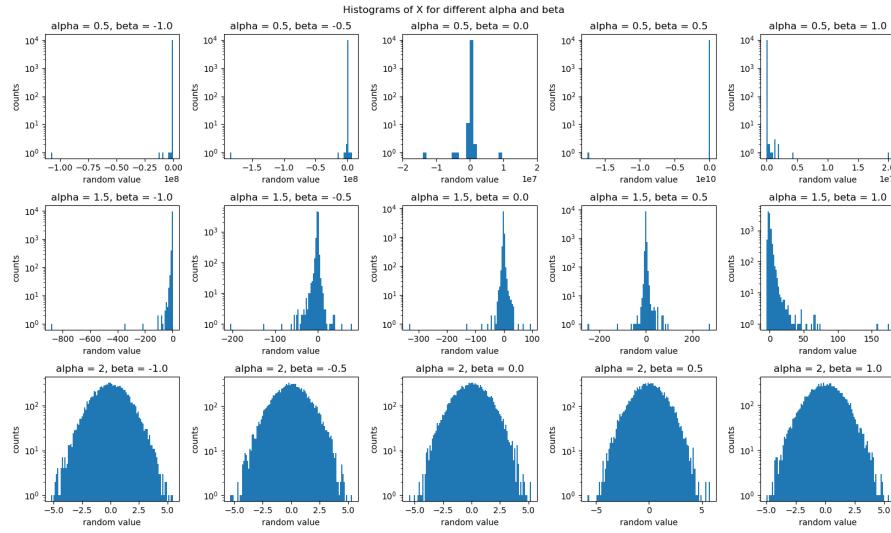


Figure 12: Histogram of  $\alpha$ -stable random variables with different  $\alpha$  and  $\beta$

From Figure 12, we can see that  $\alpha$  determines the extent of the data with extreme values. A small value of  $\alpha$  gives a distribution that has more extreme values, while a large value of  $\alpha$  gives a distribution that is more concentrated around the mean. When  $\alpha = 2$ , we can see that the distribution has become a standard Gaussian distribution.

The value of  $\beta$  determines the skewness of the distribution. A positive  $\beta$  skews the distribution to the right, and a negative  $\beta$  skews the distribution to the left.

For the case that  $\beta = 0$ , the tail probability of  $\alpha$ -stable random variables is plotted in Figure 13.

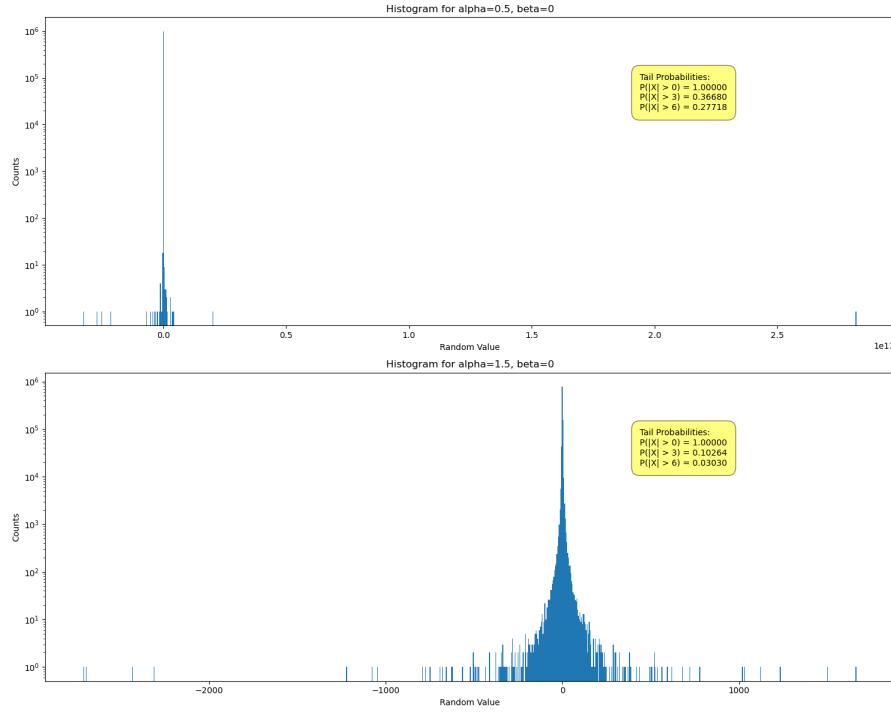


Figure 13: Tail Probability of  $\alpha$ -stable random variables with different  $\alpha$

To compare with the standard Gaussian distribution,  $\mathcal{N}(0, 1)$ , that has a tail probability of:

$$\begin{aligned} P(|X| > 0) &= 1 \\ P(|X| > 3) &\approx 2.7 \times 10^{-3} \\ P(|X| > 5) &\approx 5.7330 \times 10^{-7} \end{aligned}$$

It is clear that stable distributions generally have heavier tails than the Gaussian distribution for  $\beta = 0$ .

## 6 Conclusion

## References

- [1] Cambridge University Engineering Department. *Random Variables and Random Number Generation Lab Sheet*. 2025.
- [2] S Sinharay. *Discrete Probability Distributions*. ETS, Princeton, NJ, USA, 2010