DS 298 - Random Variates in Computation
Homework 1

N Surya (22938)

February 11, 2024

Question 1

Explanation:

From this problem, the nature of KS statistic and its independence from the nature of the distribution itself as the sample size increases is explored. All relevant plots are provided after the explanation.

- 500 trials per sample size was adopted for stable results. Mean confusion matrix is used to arrive at conclusions in order to avoid incorrect inferences. The number 500 was obtained by experimentation.
- It can be seen from the KS log-plot that after averaging over 500 trials, all KS statistics converge with the same slope.
- This conclusion was confirmed using a linear interpolation.

Cut-off for KS:

As observed from the confusion matrices, KS value (D_n) is independent of the distribution as the sample size increases. So, a theoretical analysis is required to establish this bound. This conclusion makes sense from the perspective of Central Limit Theorem also.

$$D_n \to 0; \quad n \to \infty$$

The two samples are drawn from the same distribution if for some $c > 0$,

$$D_{m,n} \leq c\sqrt{\frac{m+n}{mn}} = c\sqrt{\frac{AM}{GM}}$$

The variance of the sample distribution which is a measure of square of the distance varies as $\frac{\sigma^2}{n}$ where n is the number of samples. Hence, the K-S statistic should also vary as $\frac{c\sqrt{2}\sigma}{\sqrt{n}}$.

$$D_n \leq \frac{c\sqrt{2}\sigma}{\sqrt{n}}$$

Plots:

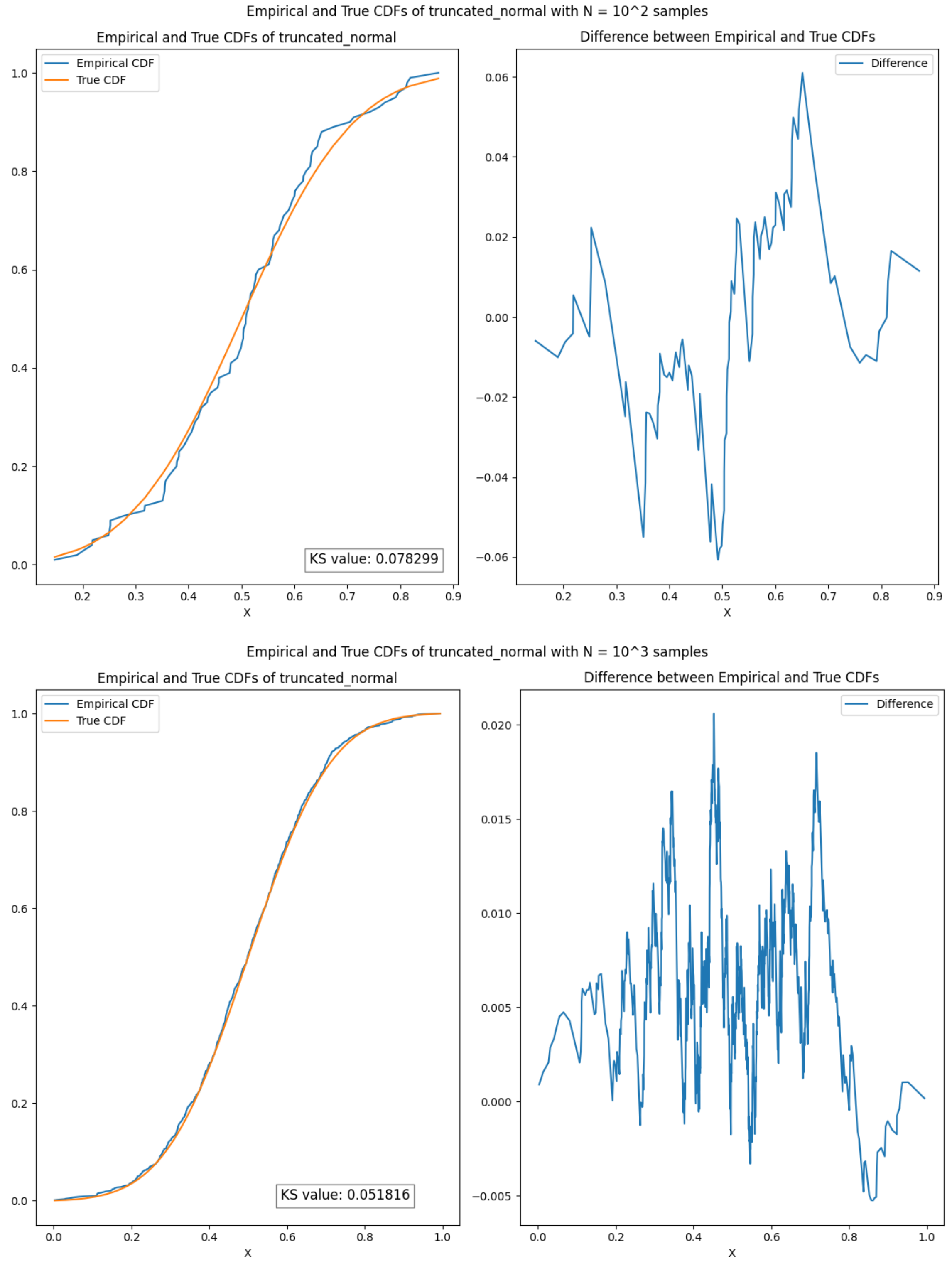


Figure 1: Truncated Normal Distributions

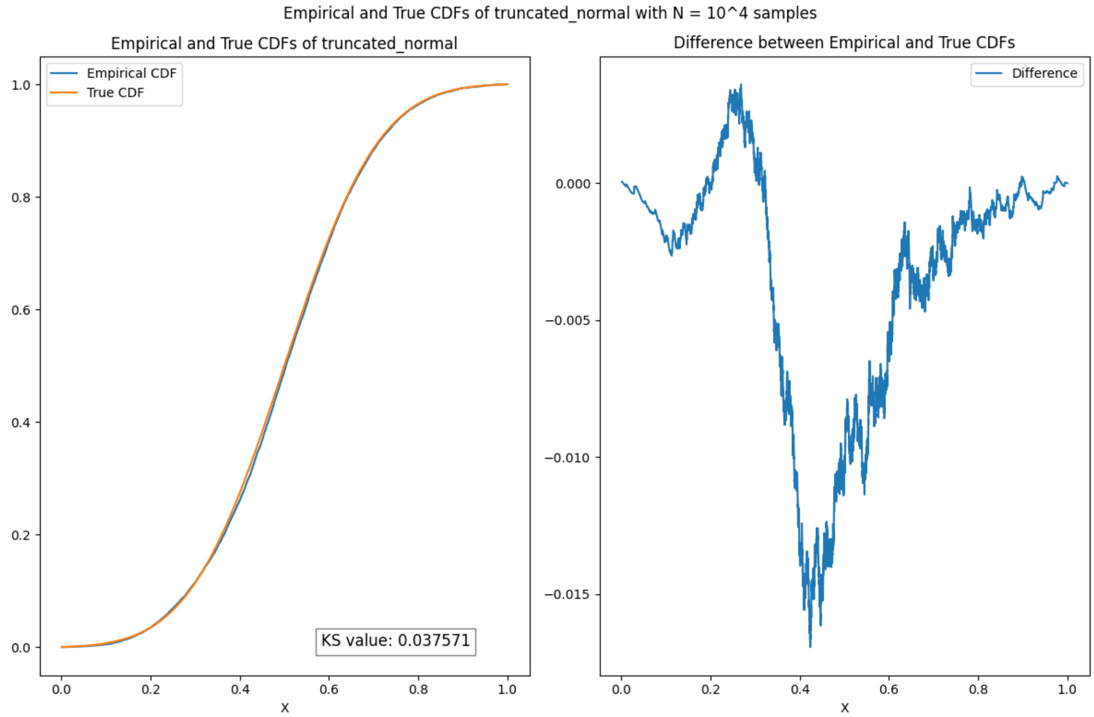


Figure 2: Truncated Normal Distribution

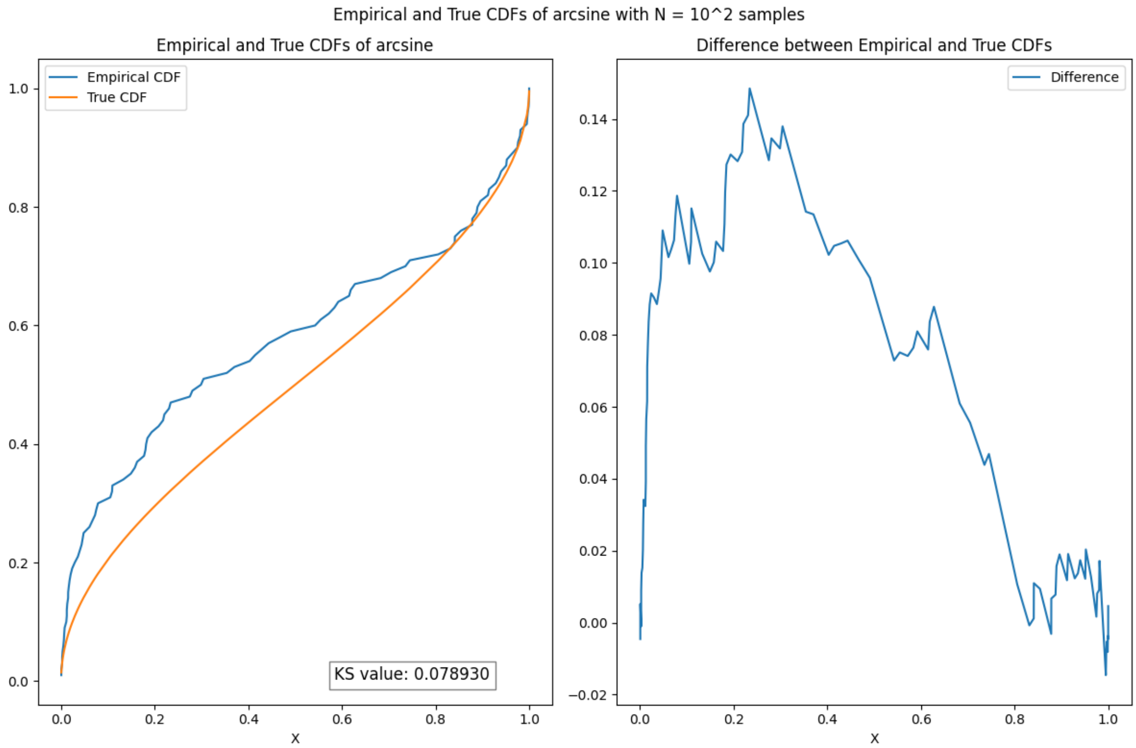


Figure 3: Arcsine Distribution

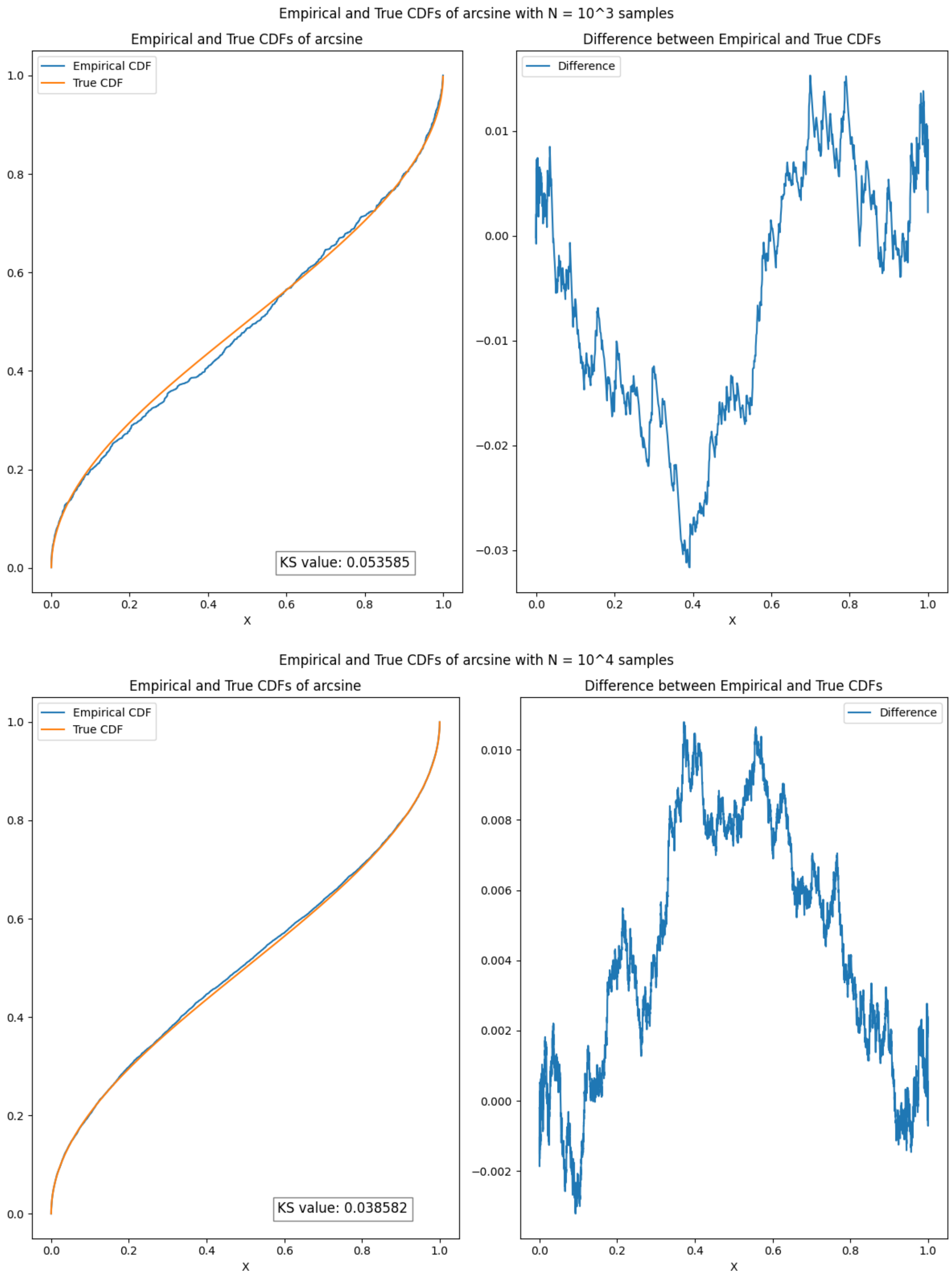


Figure 4: Arcsine Distributions

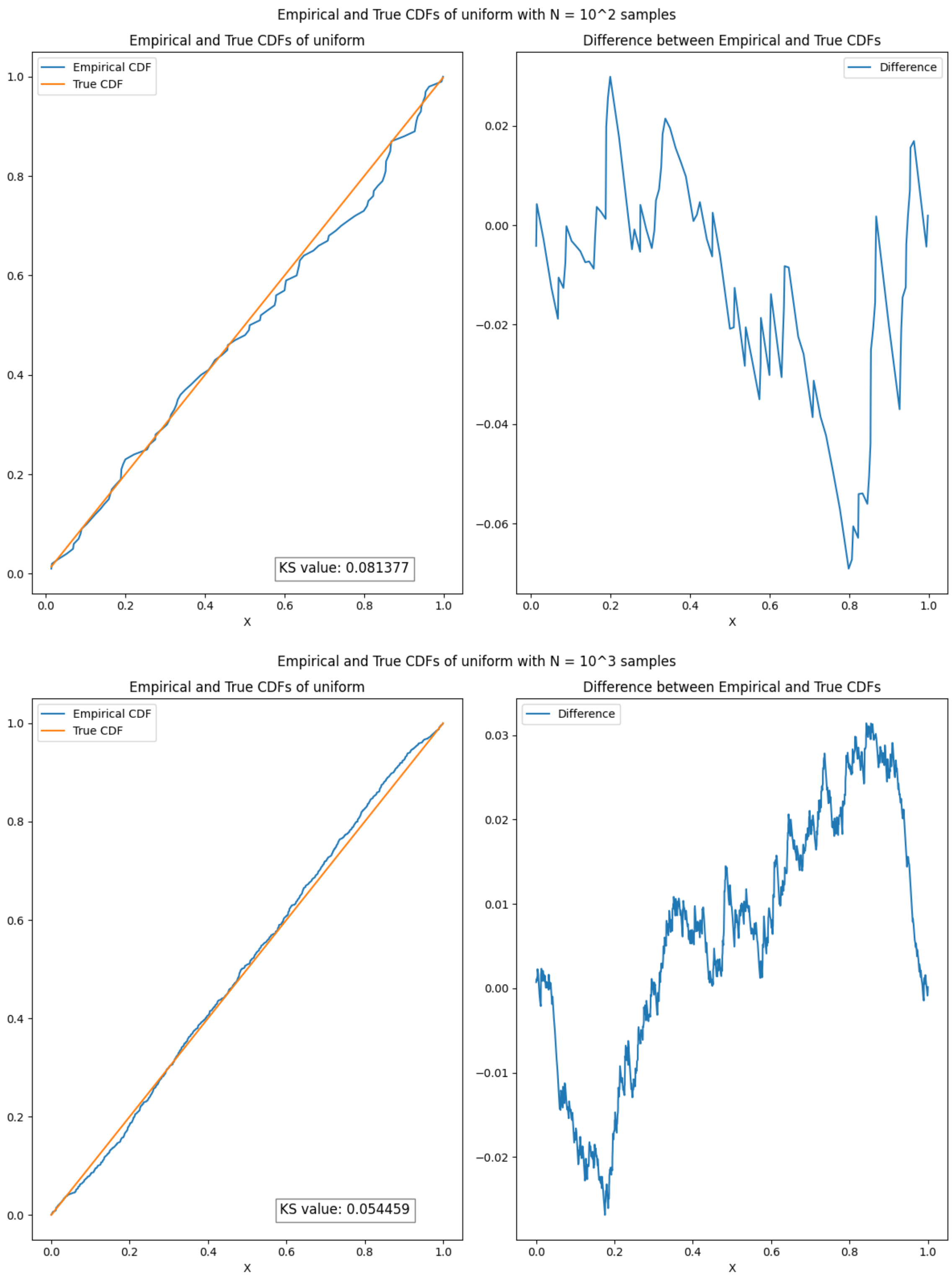


Figure 5: Uniform Distributions

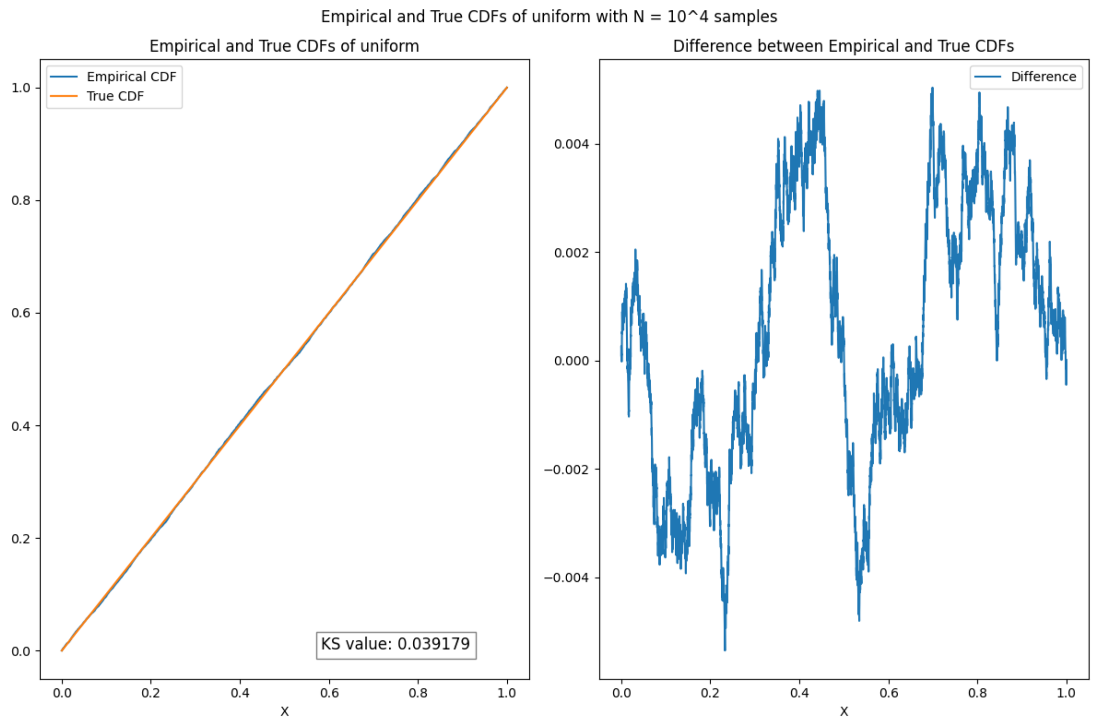


Figure 6: Uniform Distribution

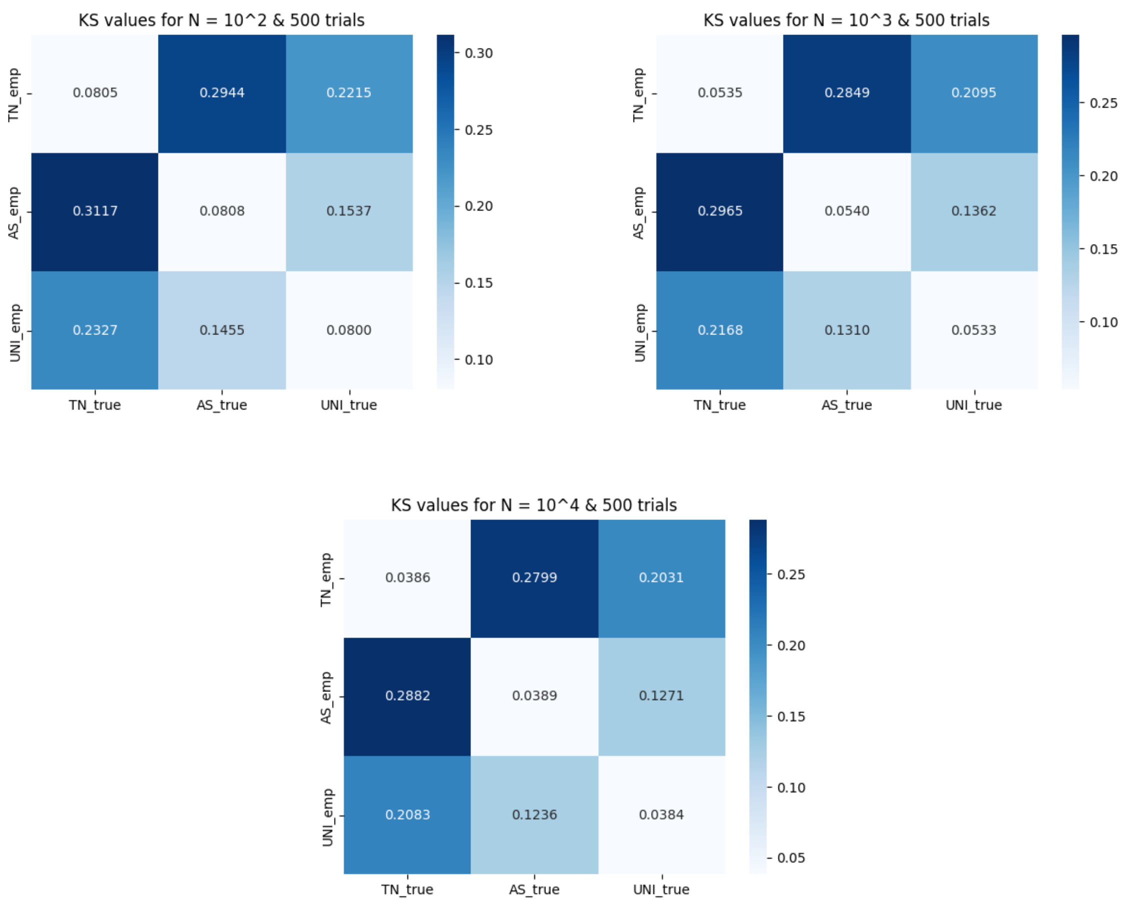


Figure 7: Confusion Matrices

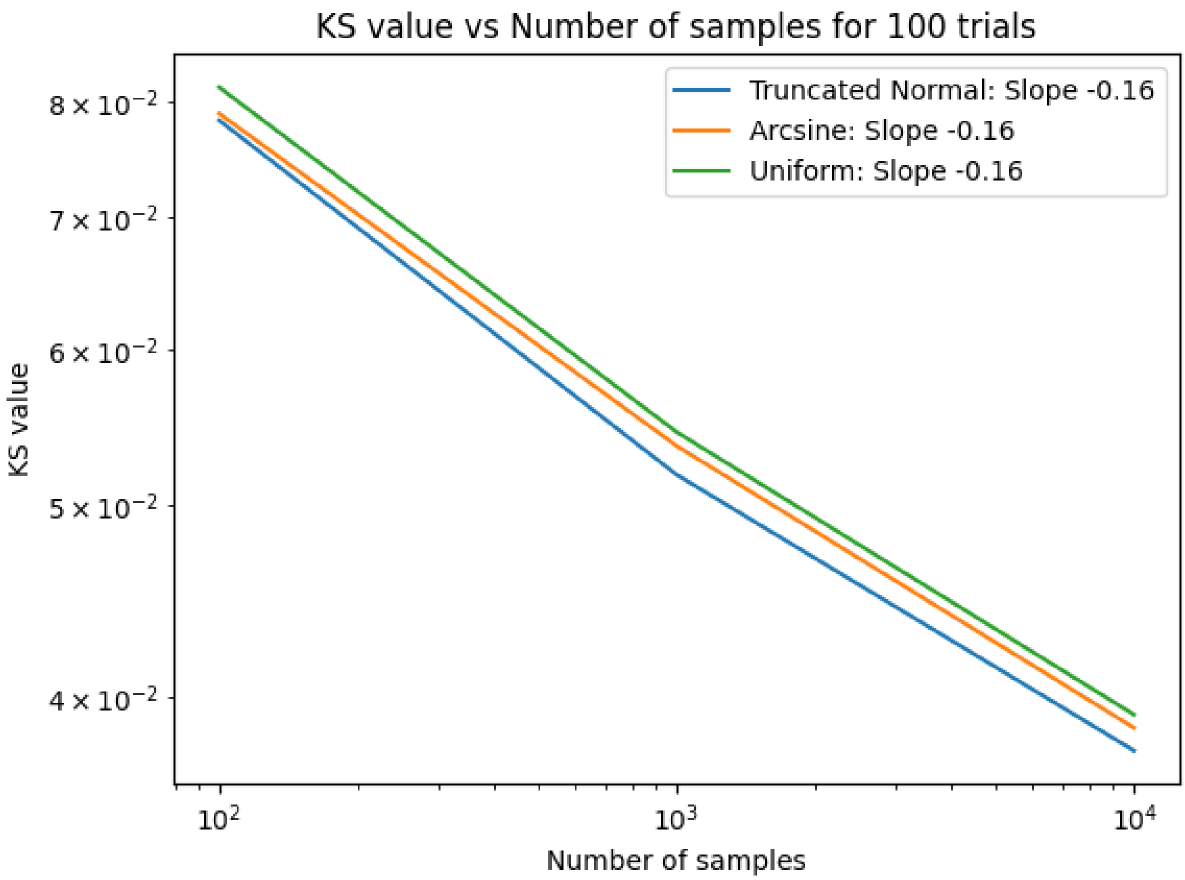


Figure 8: KS Values

Question 2

Given:

Yield (Y) in tonnes/acre is distributed as $\mathcal{N}(3, 1)$.

The speculative local pricing model in price/tonne is a exponentially decreasing probability distribution of the local yield $X = ae^{-\frac{y}{b}}$

To get the values of a and b, the given initial condition and mean condition can be used.

Given,

$$X = ae^{-\frac{Y}{b}}$$

for zero local yield, $Y = 0$

$$\implies X = a = 5000$$

$$\boxed{\text{a=5000} \quad \text{Rs./tonne}}$$

For mean local yield, $X = 5000/e \quad Rs/tonne$

$$\implies \frac{5000}{e} = 5000e^{-\frac{3}{b}}$$

$$\implies ln(1) = \frac{-3}{b} + 1$$

$$\boxed{\text{b=3} \quad \text{acre/tonne}}$$

Transformation of Variables:

$$f_X(x) = f_Y(y) \cdot |g'(x)|$$

Where:

$$X = ae^{-\frac{Y}{b}} \implies \frac{X}{a} = e^{-\frac{Y}{b}}$$

$$\implies ln(\frac{X}{a}) = -\frac{Y}{b} \implies y = -b.ln(\frac{x}{a})$$

$$g'(x) = -\frac{b}{(x/a)} \implies -\frac{ba}{x}\frac{1}{a} \implies \left|-\frac{b}{x}\right|$$

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

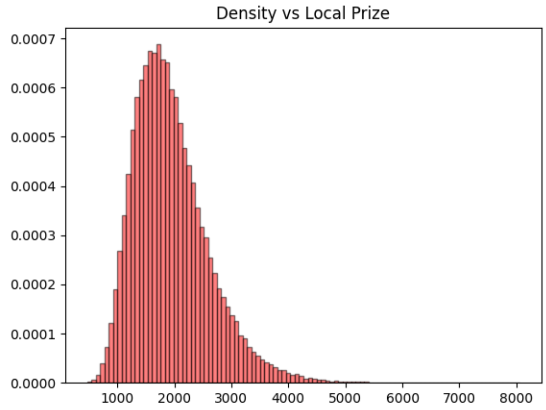
$$f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{(y-3)^2}{2}} \cdot \left|-\frac{b}{x}\right|$$

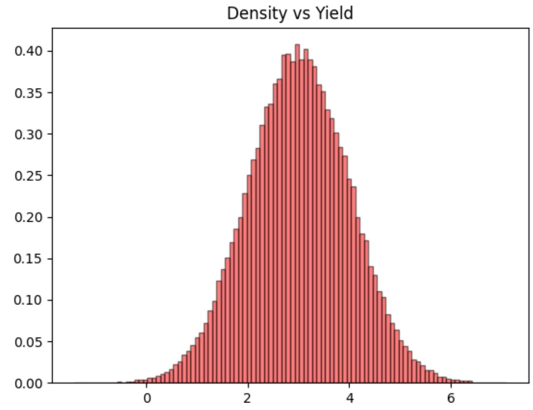
where $\mu = 3$ and $\sigma = 1$

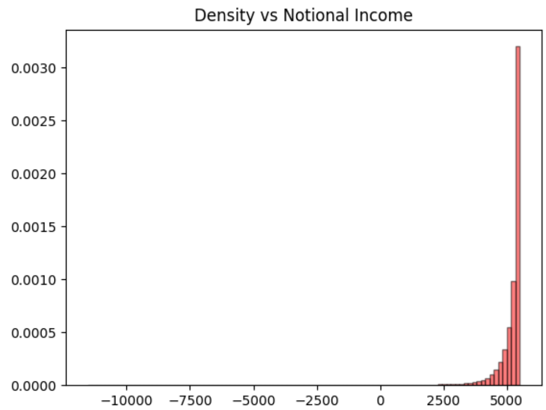
$$f_X(x) = \frac{1}{2\pi} e^{-\frac{\left(-3\ln\left(\frac{x}{5000}\right) - 3\right)^2}{2}} \cdot \left|\frac{-3}{x}\right|$$

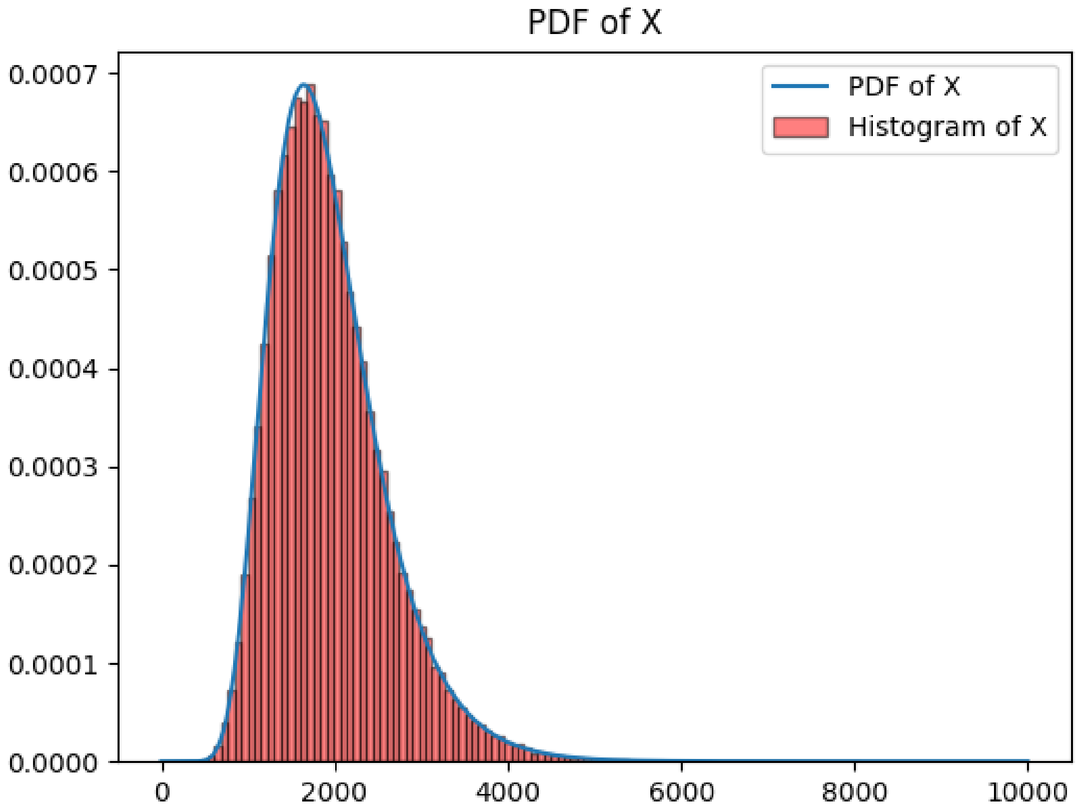
Conclusions and Plots:

- The mean for Notional Income is 5195.128610.
- The histogram for Z = XY agrees with the transformed PDFs.









PDF of X
PDF of X
Histogram of X
0.0007
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
0.0000
0
2000
4000
6000
8000
10000