Part A: With inverse transformation.

In the theorem, it is indicated that if X is a continuous random variable with CDF F(x). Define a random variable U = F(x) where we have U (Uniform(0, 1)).

To generate variable X with a given CDF, first we define U = F(X).

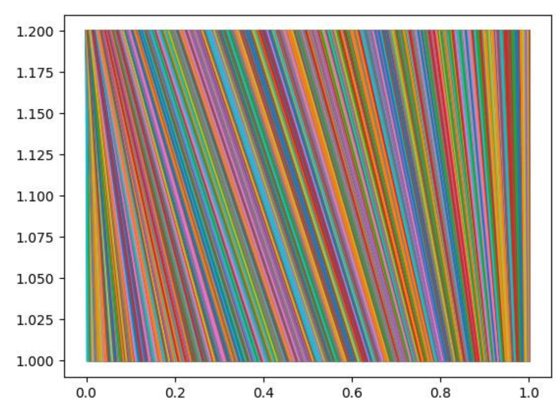
After that we can generate X from a generated uniform standard variable U with the inverse transformation.

$X = F^{-1}(U)$

PROOF: First, we notice that $0 \leq F(x) \leq 1$ for all x, therefore, values of U lie in $[0,1]$. Second, for any $u \in [0,1]$, find the cdf of U,

$$\begin{aligned} F_U(u) &= \boldsymbol{P}\{U \leq u\} \\ &= \boldsymbol{P}\{F_X(X) \leq u\} \\ &= \boldsymbol{P}\left\{X \leq F_X^{-1}(u)\right\} && \text{(solve the inequality for } X\text{)} \\ &= F_X(F_X^{-1}(u)) && \text{(by definition of cdf)} \\ &= u && (F_X \text{ and } F_X^{-1} \text{ cancel}) \end{aligned}$$

Figure 1: *x versus u relationship*



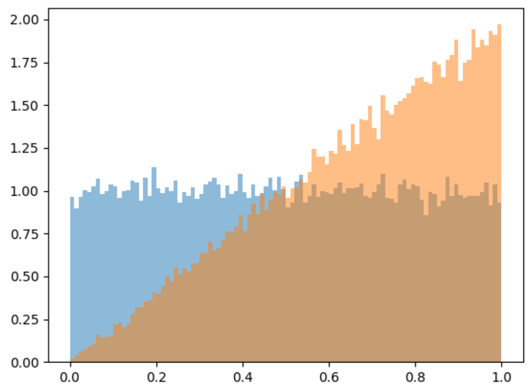
In **figure 1**, each different line shows the random variable x (at y-axis = 1.000) and its corresponding standard uniform random variable u (at y-axis == 1.200). It is clearly seen that the difference between x and u have little around the end of the intervals, this is caused by the inverse transformation between the cumulative distribution function (which the generated x will have) and u.

$F(x) = x^2$ for $0 \leq x \leq 1$

And, using inverse transformation: **$F(x) = x^2 = u$**

$F(0) = 0$, $F(0.2) = 0.04$, $F(0.5) = 0.25$, … , $F(1) = 1$

Figure 2:



In **figure 2**, the blue histogram shows a continuous line, which shows us the standard uniform random variable "u", which means as a whole it shows standard uniform distribution function of u, which is f(u). And the brown histogram shows the probability distribution function of x, which is f(x):

f(x) is reached by taking the derivative of F(x):

since **$F(x) = x^2$ → $f(x) = 2x$**

As the number of trials increase, the histograms become have straight line.

Figure 3:

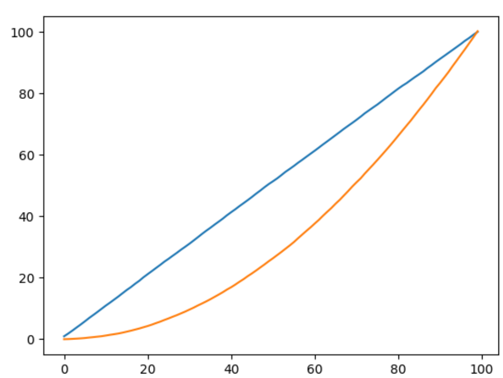


Figure 3 shows two cumulative distribution functions. Blue line stands for the cumulative distribution function of the standard uniform distribution. The yellow line stands for the cumulative distribution function of the x, which is F(x) = x^2 for $0 \le x \le 1$.

Calculations:

We can reach the theoretical results of the mean and variance by the following calculations.

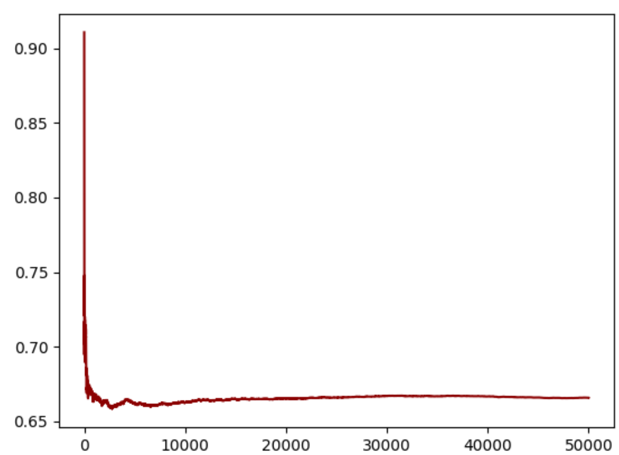
Expected value of x: **E[x]** = $\int_0^1 xf(x)\,dx$, so $\int_0^1 x(2x)\,dx \rightarrow 2\int_0^1 x^2\,dx \rightarrow 2\left[\frac{x^3}{3}\right]_0^1$

→ 2 (1/3 - 0) → 2/3 ≈ 0.666667

Variance of x: **Var[x]** = $\int_0^1 x^2 f(x)\,dx - [\int_0^1 xf(x)\,dx]^2$

$\rightarrow 2\int_0^1 x^3\,dx - \frac{4}{9} \rightarrow 2\left[\frac{x^4}{4}\right]_0^1 - \frac{4}{9}$ **≈ 0.0555556**

Figure 4:



It is found from the calculation that, the mean is 0.666667. It is clearly seen that at the beginning of the experiment since there were few sample, the line is fluctuating. With more data, the line gets smooth, and approach the theoretical result of the mean which is 0.666667.

Let see more clearly:

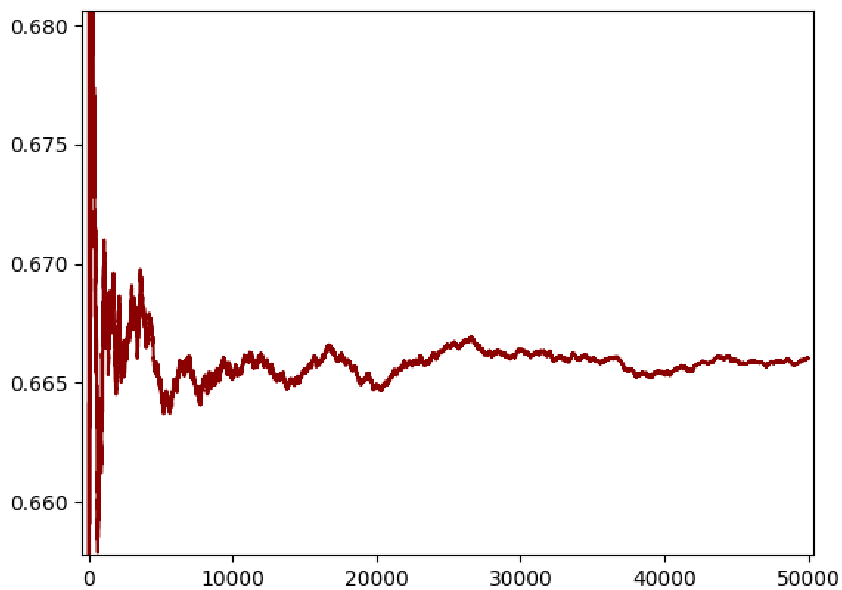
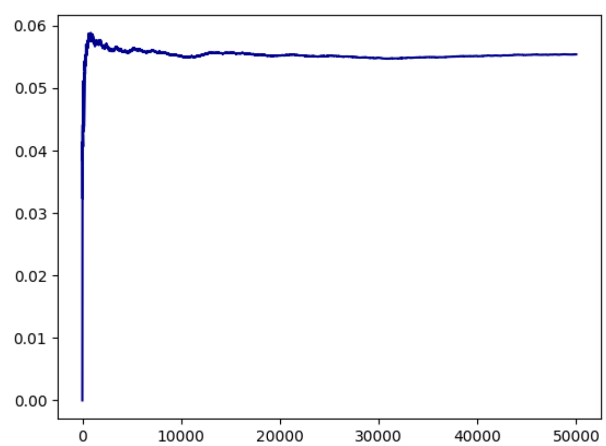


Figure 5:



The theoretical calculations give us 0.0555556. The line plot shows that, at the beginning of the experiment (when there're few samples), the line fluctuating. Gradually, the line becomes smooths and approaches its theoretical result.

Let see more clearly:

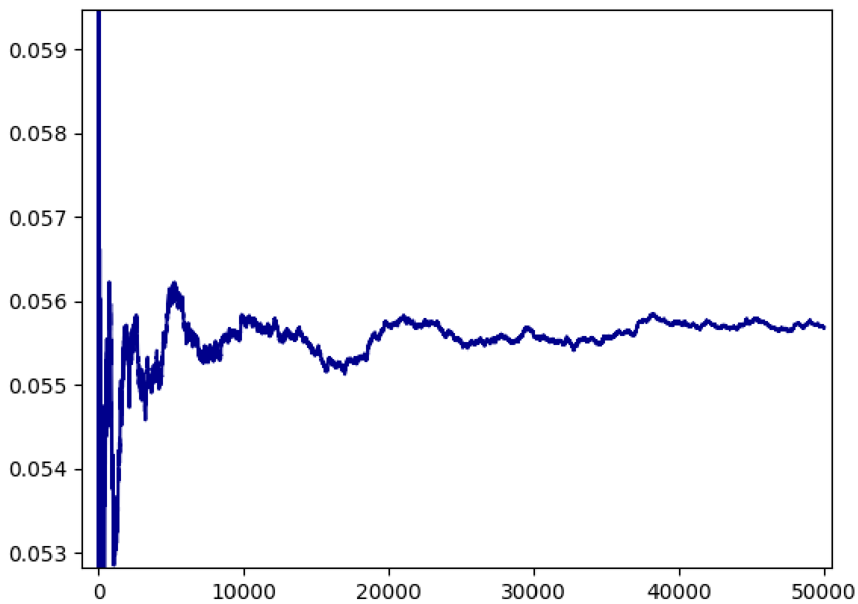


Figure 4 and **figure 5** shows the updated mean and variance of the X.

Part B: With rejection method.

When F(x) has a complicated form but a density f(x) is available, random variables with this density can be generated by rejection method.

Theorem 3 *Let a pair (X, Y) have Uniform distribution over the region*

$$A = \{(x, y) \mid 0 \leq y \leq f(x)\}$$

for some density function f. Then f is the density of X.

Algorithm:

1. Find such numbers a, b, and c that $0 \leq f(x) \leq c$ for $a \leq x \leq b$. The bounding box stretches along the x-axis from a to b and along the y-axis from 0 to c.
2. Obtain Standard Uniform random variables U and V from a random number generator or a table of random numbers.
3. Define $X = a + (b - a)U$ and $Y = cV$. Then X has Uniform(a, b) distribution, Y is Uniform$(0, c)$, and the point (X, Y) is Uniformly distributed in the bounding box.

4. If $Y > f(X)$, reject the point and return to step 2. If $Y \leq f(X)$, then X is the desired random variable having the density $f(x)$.

Figure 6:

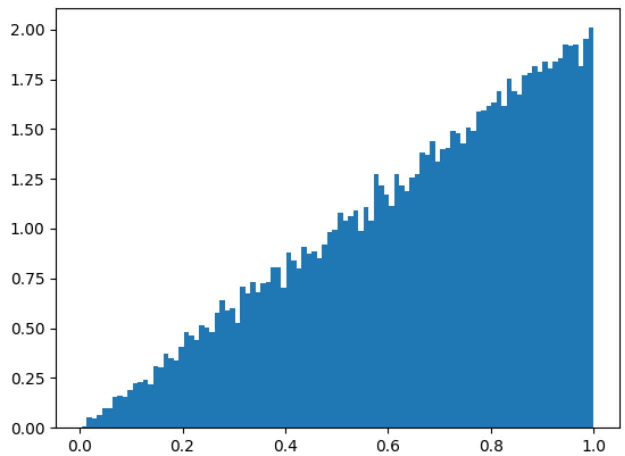


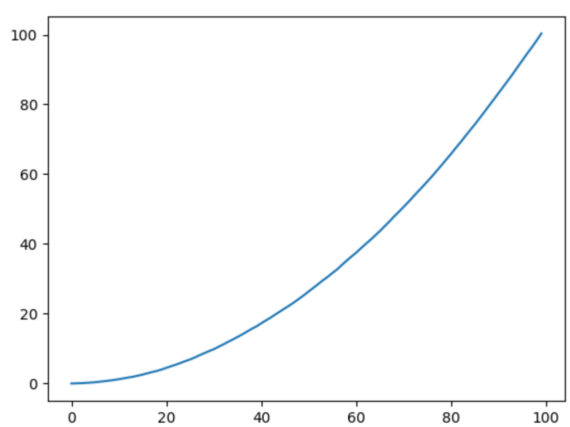
Figure 6 indicates the probability density function of x. It has a linear like shape because of the following calculations:

f(x) = F'(x) We know that **F(x) = x^2** So, **f(x) = F'(x) = 2 * x**

We can see the correspondence of the function with the plot:

f(0) = 0, f(0.5) = 1, f(1) = 2 So this plot shows us the f(x), which is 2*x for $0 \leq x \leq 1$.

Figure 7:

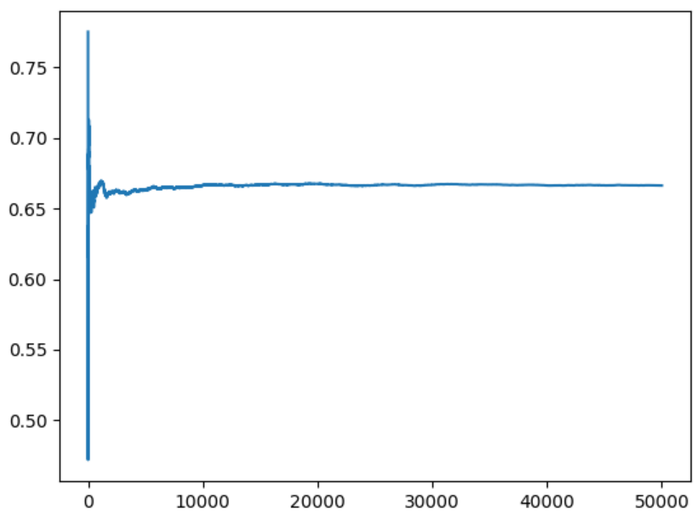


It is clearly seen that the graph corresponds to the cumulative distribution function (CDF) which is

$F(x) = x^2$ for $0 \le x \le 1$.

We can see that in **figure 3**, the brown line shows the same line as **figure 7**, since they both indicate the same cumulative distribution function.

Figure 8: Average changes throughout the experiment.



At the beginning of the process, with the Rejection method, since the sample size is low, the plot (which shows the mean) is fluctuated.

As the sample size is increasing, the line approaches the theoretical value that we have found in the **Calculations** section.

As we found the mean is 0.666667, the line approaches 0.666667 as more sample goes into experiment.

Let look the sandwiched area more closely:

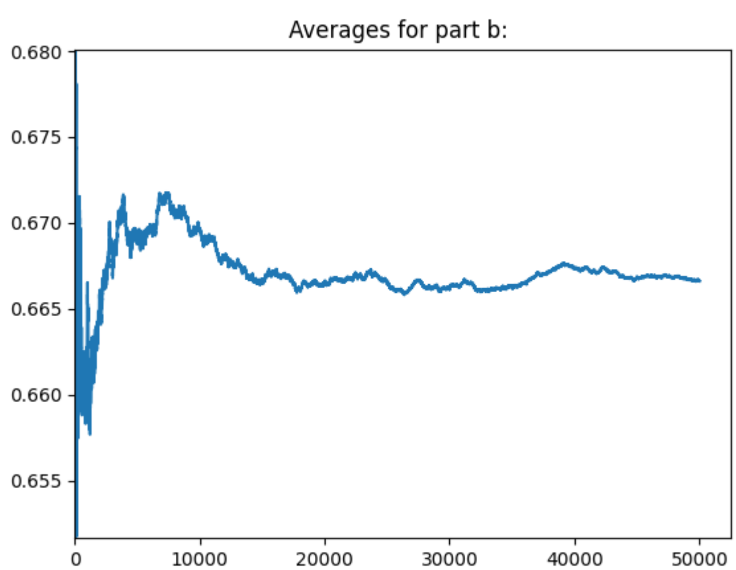
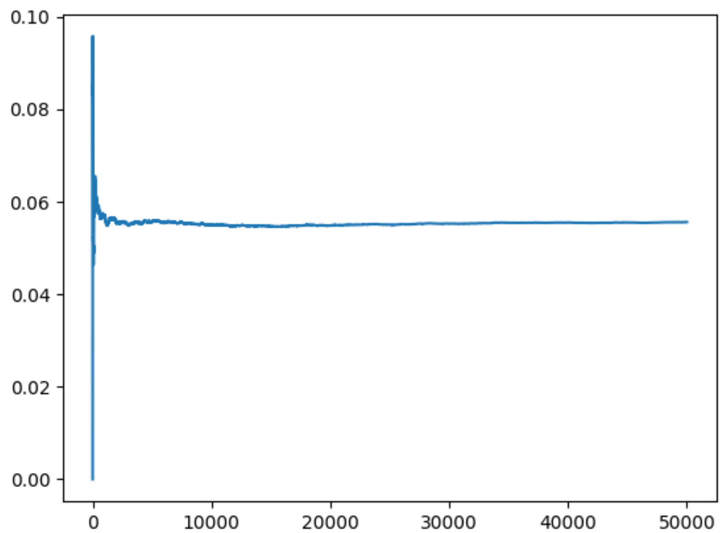


Figure 9: Variance changes throughout the experiment.



As the process begins the variance of the sample is fluctuated as in the **part a figure 5**. They have nearly the same plot. They both approaches the theoretical result of the variance while more and more samples are get into the experiment.

Results:

By creating sample trials, we want to see the methods that we learned are working. With inverse transformation method or rejection method we approaches the theoretical results that we have reached: So if we have pdf, we can use the rejection method; if we have cdf (and that cdf can be used in the inverse transformation), then we can use the inverse transformation method. And from the theoretical calculations, we can see that the method that we used are correct.

CENG222 Homework 3 / Report Name: Çağın Tunç
ID: 280201090

Let see our approximations: