MA323 Lab 5

Ashish Kumar Barnawal - 180123006

October 7, 2020

Question 1

(a)

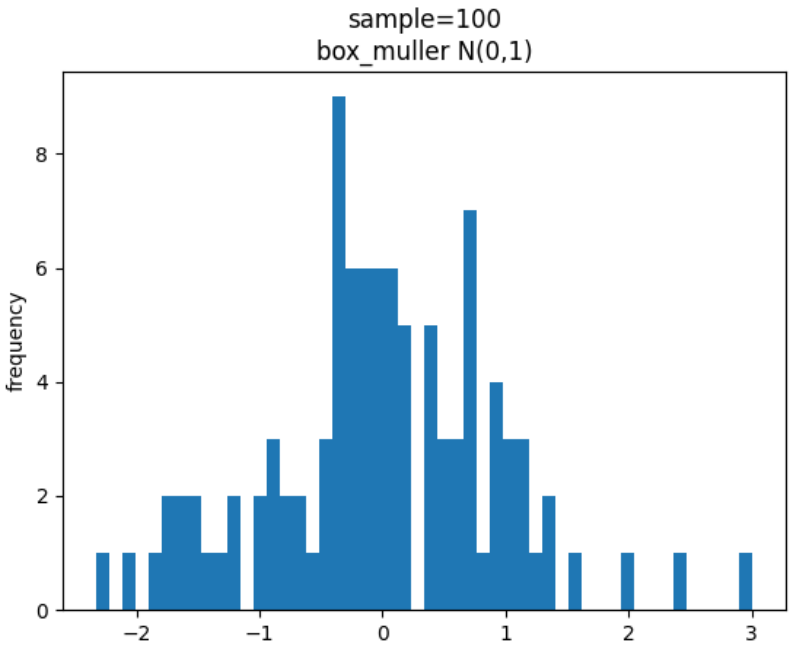
Table 1: Box-Muller

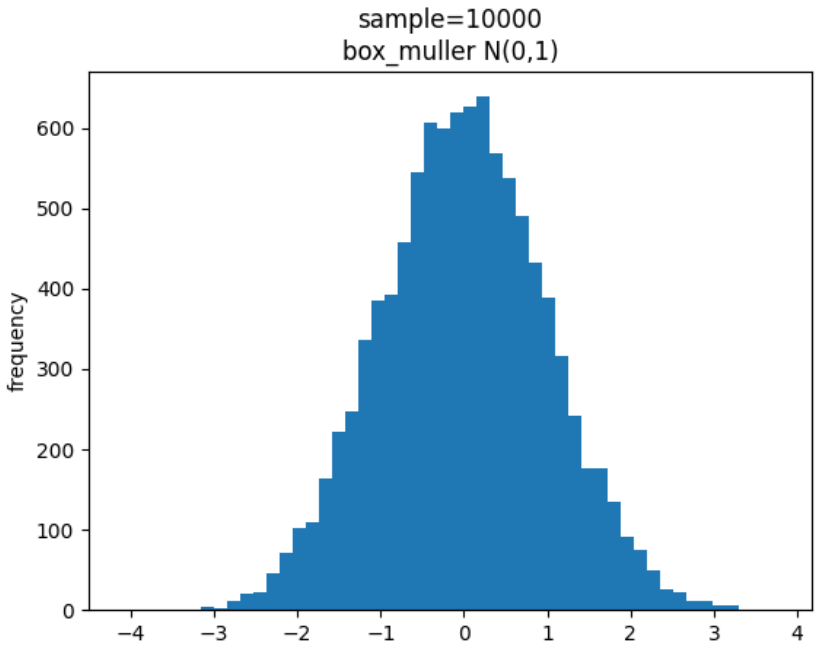
sample size	mean	variance
100	0.0565	0.82513
10000	-0.0151	1.00008

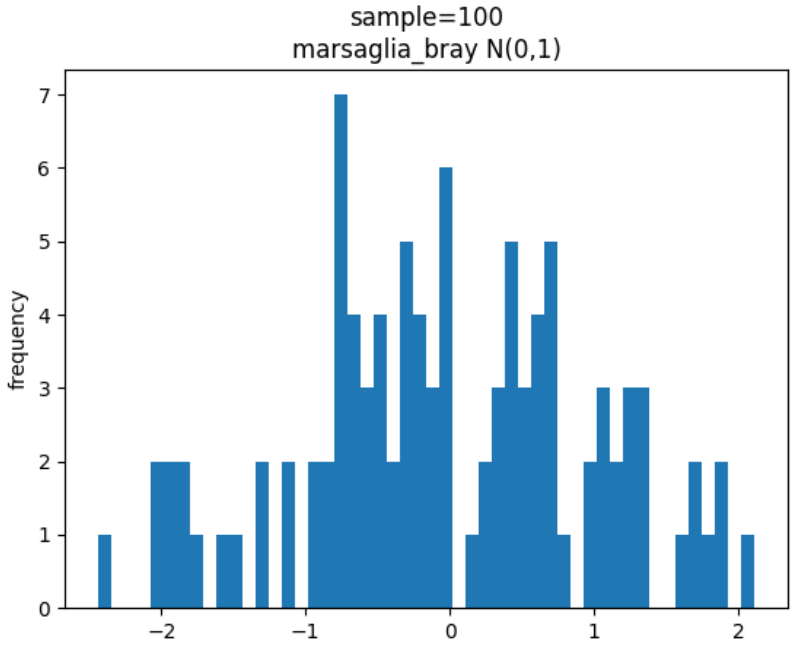
Table 2: Marsaglia and Bray

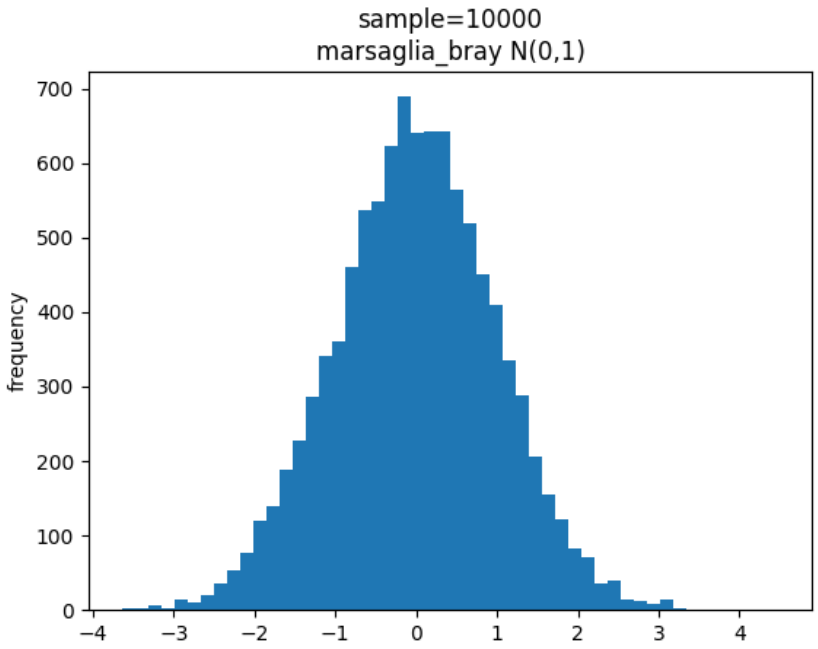
sample size	mean	variance
100	-0.0389	0.93992
10000	-0.0054	1.01210

(b)




sample=10000
box_muller N(0,1)
frequency
600
500
400
300
200
100
0
−4
−3
−2
−1
0
1
2
3
4


sample=100
marsaglia_bray N(0,1)
frequency
7
6
5
4
3
2
1
0
−2
−1
0
1
2

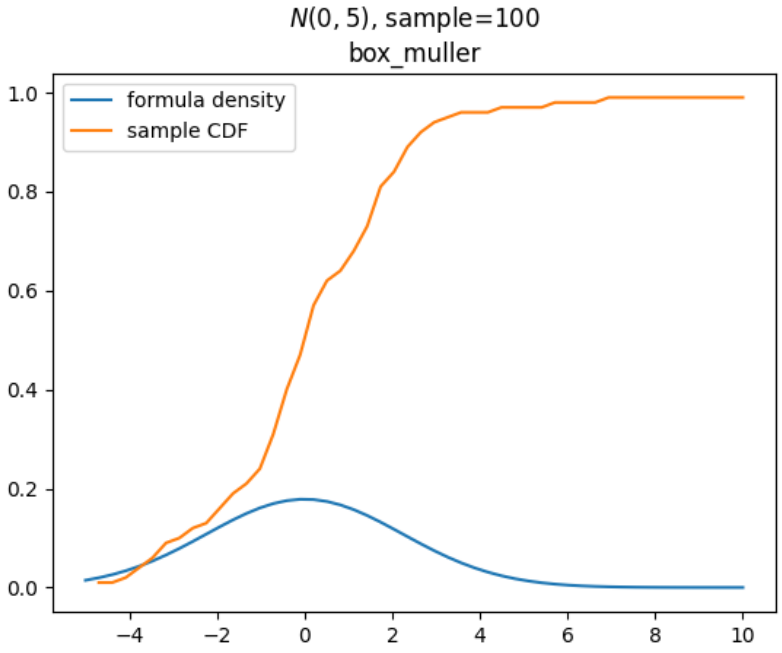


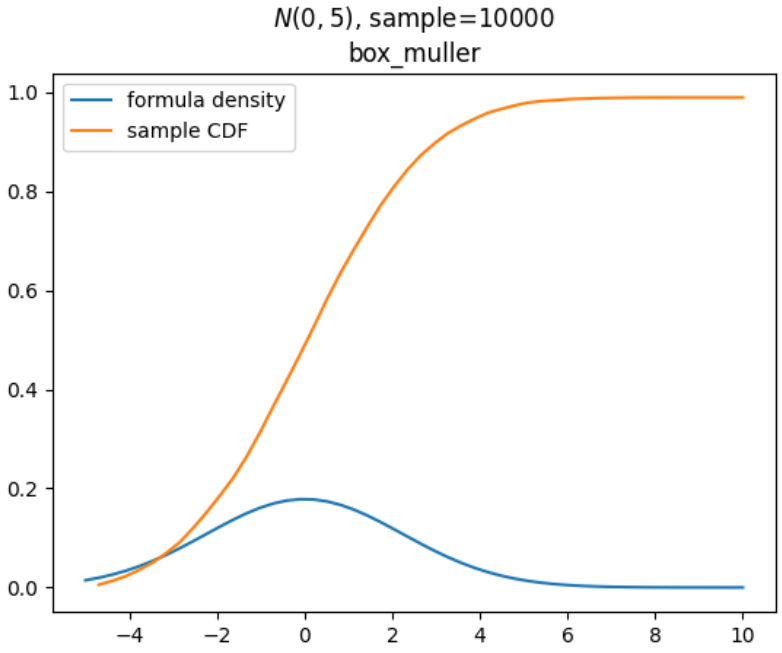
(c)

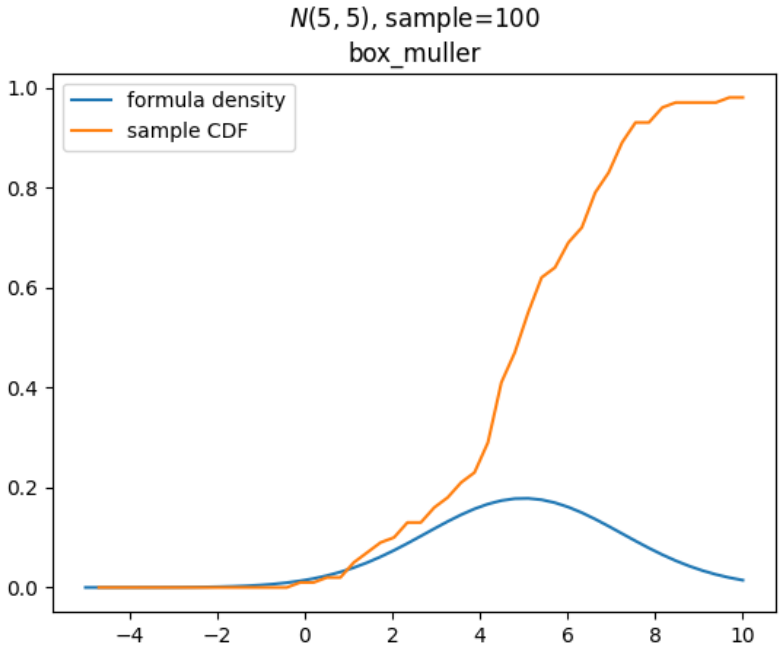
We can see that the sample CDF increases more rapidly when the formula density is higher. This is because the theoritically CDF at point $x = t$ is

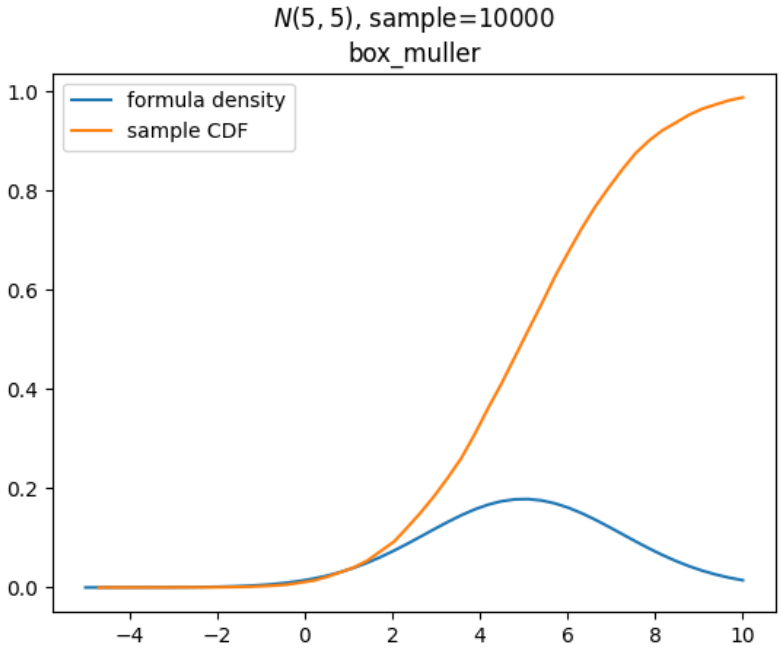
$$P(X \leq t) = F(t) = \int_{-\infty}^{t} f(x)dx$$

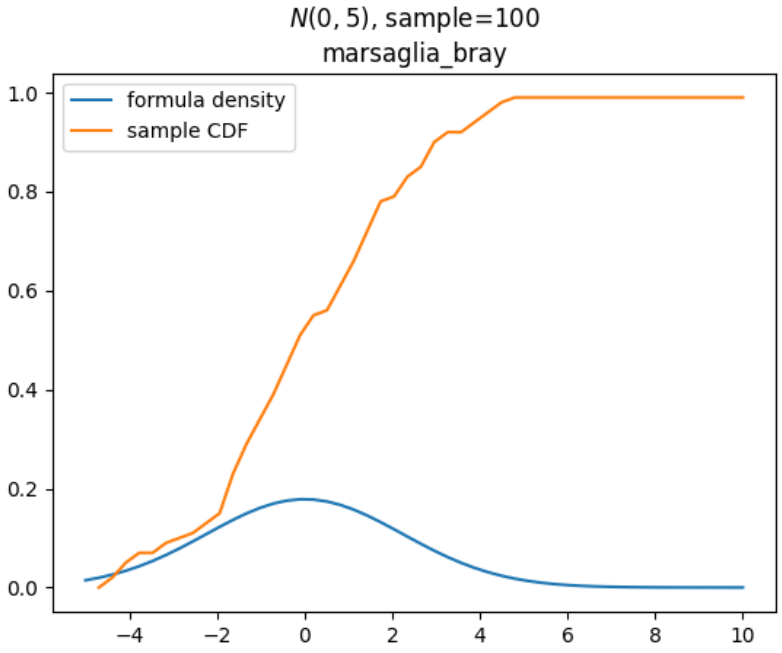
where $f(x)$ is the density function of $N(0, 1)$ distribution

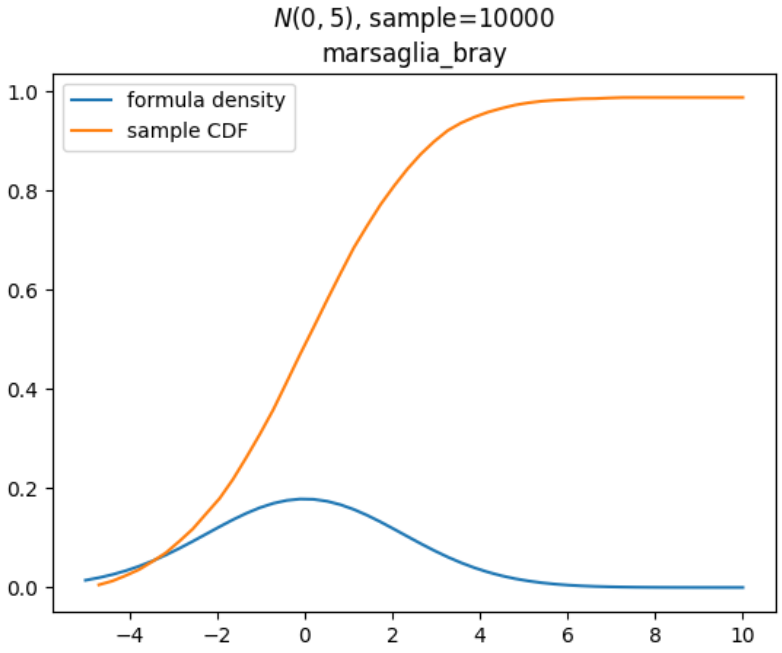

$N(0, 5)$, sample=100
box_muller
formula density
sample CDF
1.0
0.8
0.6
0.4
0.2
0.0
−4
−2
0
2
4
6
8
10

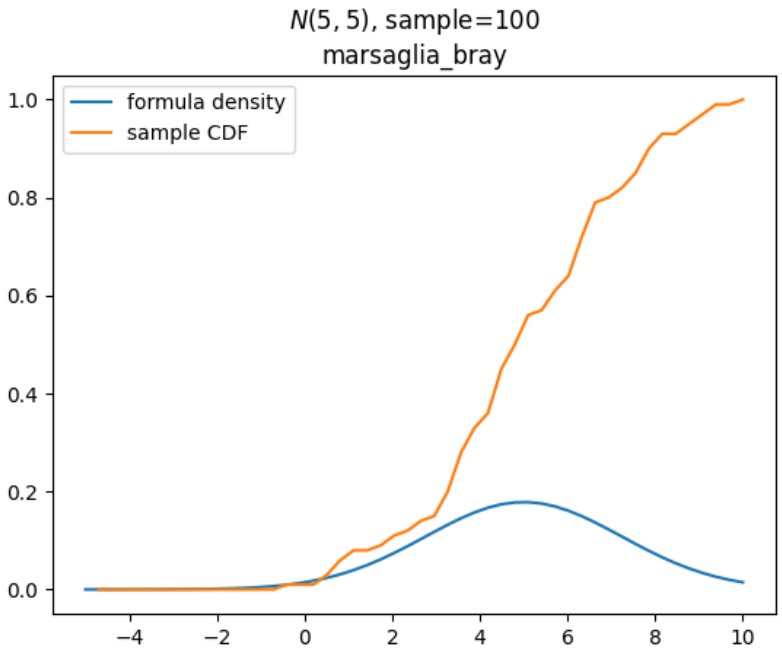

$N(0, 5)$, sample=10000
box_muller
formula density
sample CDF
1.0
0.8
0.6
0.4
0.2
0.0
−4
−2
0
2
4
6
8
10

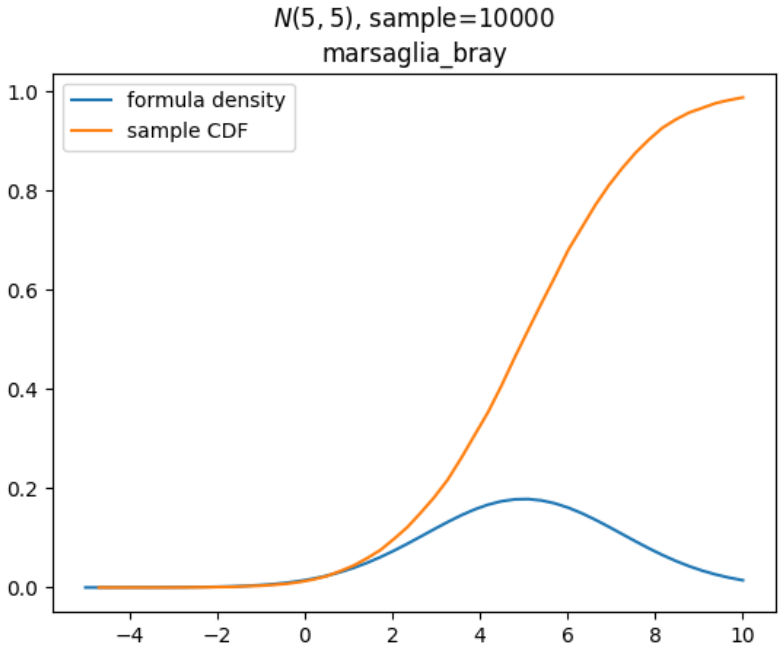

$N(5, 5)$, sample=100
box_muller
formula density
sample CDF
1.0
0.8
0.6
0.4
0.2
0.0
−4
−2
0
2
4
6
8
10


$N(5, 5)$, sample=10000
box_muller
formula density
sample CDF
1.0
0.8
0.6
0.4
0.2
0.0
−4
−2
0
2
4
6
8
10


$N(0, 5)$, sample=100
marsaglia_bray
formula density
sample CDF
1.0
0.8
0.6
0.4
0.2
0.0
−4
−2
0
2
4
6
8
10


$N(0, 5)$, sample=10000
marsaglia_bray
formula density
sample CDF
1.0
0.8
0.6
0.4
0.2
0.0
−4
−2
0
2
4
6
8
10


$N(5, 5)$, sample=100
marsaglia_bray
formula density
sample CDF
1.0
0.8
0.6
0.4
0.2
0.0
−4
−2
0
2
4
6
8
10


$N(5, 5)$, sample=10000
marsaglia_bray
formula density
sample CDF
1.0
0.8
0.6
0.4
0.2
0.0
−4
−2
0
2
4
6
8
10

Question 2

Although technically Marsaglia and Bray method is faster than Box-Muller, but due to the **more intepreter heavy implementation in python**, Marsaglia and Bray's method runs slower than Box-Muller.

In the Box-Muller method, the calculation of *sin* and *cos* are the steps taking the most times, but in our python code, these functions are implemented in C and are very fast. While in Marsaglia and Bray's method, we've almost all the code in pure python. Python being an interpreted language, does a lot of type checking which adds to the complexity. Thus making Marsaglia and Bray method to run slower.

Table 3: Computation times

sample size	Box-Muller	Marsaglia and Bray
100	0.0567ms	0.0906ms
10000	4.7739ms	6.6557ms

Question 3

We can see that fraction of rejects get closer to $1 - \frac{\pi}{4}$ as the sample size increases. This is due to the fact that

$$P(U_1^2 + U_2^2 \leq 1) = \frac{\text{Area of unit circle}}{\text{Area of } [-1,1] \times [-1,1]} = \frac{\pi}{2^2}$$

Hence probability of rejection is

$$P(U_1^2 + U_2^2 > 1) = 1 - \frac{\pi}{4} = 0.2146018366...$$

Table 4: Fraction rejects in Marsaglia and Bray method

samples	100	10000	100000	1000000
fraction rejected	0.152542	0.206852	0.212486	0.214857