MA 323 (2020) Monte Carlo Simulation: LAB 01
Jay Vikas Sabale
180123019

Problem I:

Tabulation of Data :

- *Case I* : $a = 6,\ b = 0,\ m = 11$

x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
0	0	0	0	0	0	0	0	0	0	0
1	6	3	7	9	10	5	8	4	2	1
2	1	6	3	7	9	10	5	8	4	2
3	7	9	10	5	8	4	2	1	6	3
4	2	1	6	3	7	9	10	5	8	4
5	8	4	2	1	6	3	7	9	10	5
6	3	7	9	10	5	8	4	2	1	6
7	9	10	5	8	4	2	1	6	3	7
8	4	2	1	6	3	7	9	10	5	8
9	10	5	8	4	2	1	6	3	7	9
10	5	8	4	2	1	6	3	7	9	10

- $Case\ II:\ a = 3,\ b = 0,\ m = 11$

x_0	x_1	x_2	x_3	x_4	x_5
0	0	0	0	0	0
1	3	9	5	4	1
2	6	7	10	8	2
3	9	5	4	1	3
4	1	3	9	5	4
5	4	1	3	9	5
6	7	10	8	2	6
7	10	8	2	6	7
8	2	6	7	10	8
9	5	4	1	3	9
10	8	2	6	7	10

How many distinct values appear before repetitions?
Ans :

- $Case\ I:\ a = 6,\ b = 0,\ m = 11$

Seeing the tabular data given above, except for the case where $x_0 = 0$, the time period is 10, also known as *full period*. After every 10 iterations the values of the linear congruence start repeating itself.
Recurrence Obtained: $x_i = x_{i+10}\ (mod\ m)$.

In the case where $x_0 = 0$; $x_i = 0\ for\ all\ i >= 0$. Hence the time period in this case is 1.

- $Case\ II:\ a = 3,\ b = 0,\ m = 11$

Seeing the tabular data given above, except for the case where $x_0 = 0$, the time period is 5. After every 5 iterations the values of the linear congruence starts repeating itself.
Recurrence Obtained: $x_i = x_{i+5}\ (mod\ m)$.

In the case where $x_0 = 0$; $x_i = 0\ for\ all\ i >= 0$. Hence the time period in this case is 1.

Which, in your opinion, are the best choices and why?
In my opinion, amongst the choices given, wherein:

- the first choice maps values from the set of Whole numbers to the set: $\{1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9,\ 10\}$ with a periodicity of 10, and

- the second choice maps the values from the set of Whole numbers to a set of 5 distinct values, depending on the choice of x_0 , again with a periodicity of 5.

In general we use such kind of mappings, as hash functions. Few characteristics of good hash functions are:

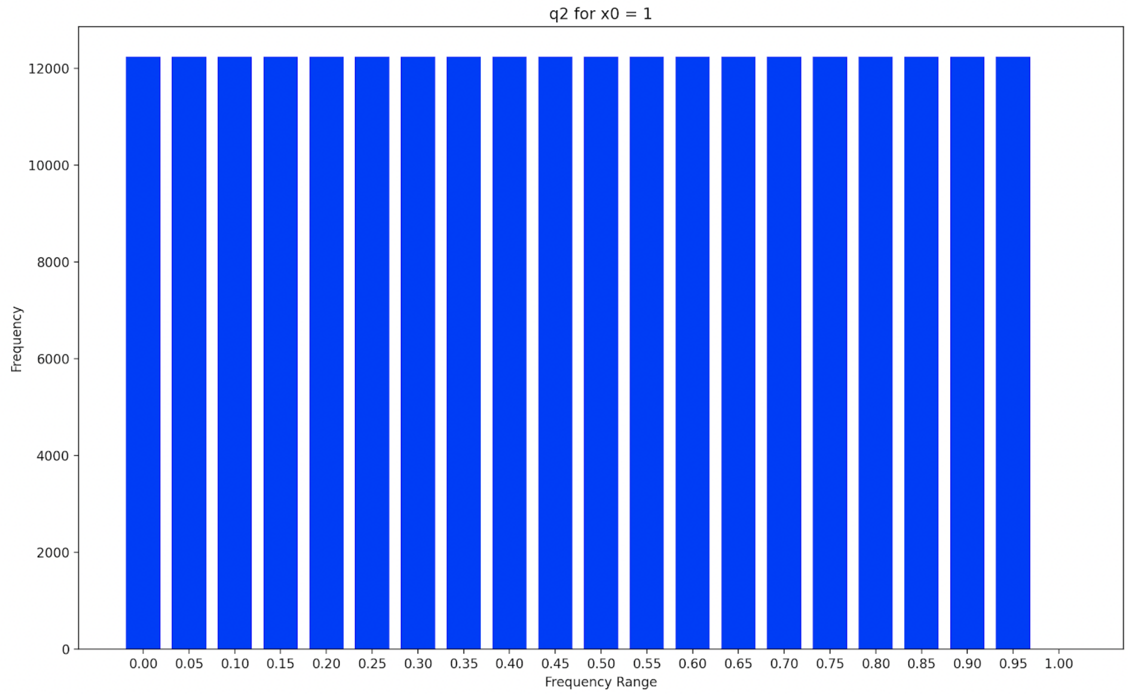
1. The hash value is fully determined by the data being hashed.
2. The hash function uses all the input data.
3. The hash function "uniformly" distributes the data across the entire set of possible hash values.
4. The hash function generates very different hash values for similar input(s).

Cleary the ideal choice of values for $a,\ b\ and\ c\ would\ be\ a = 6,\ b = 0\ and\ m = 11$. The differentiator being the third criterion mentioned above.

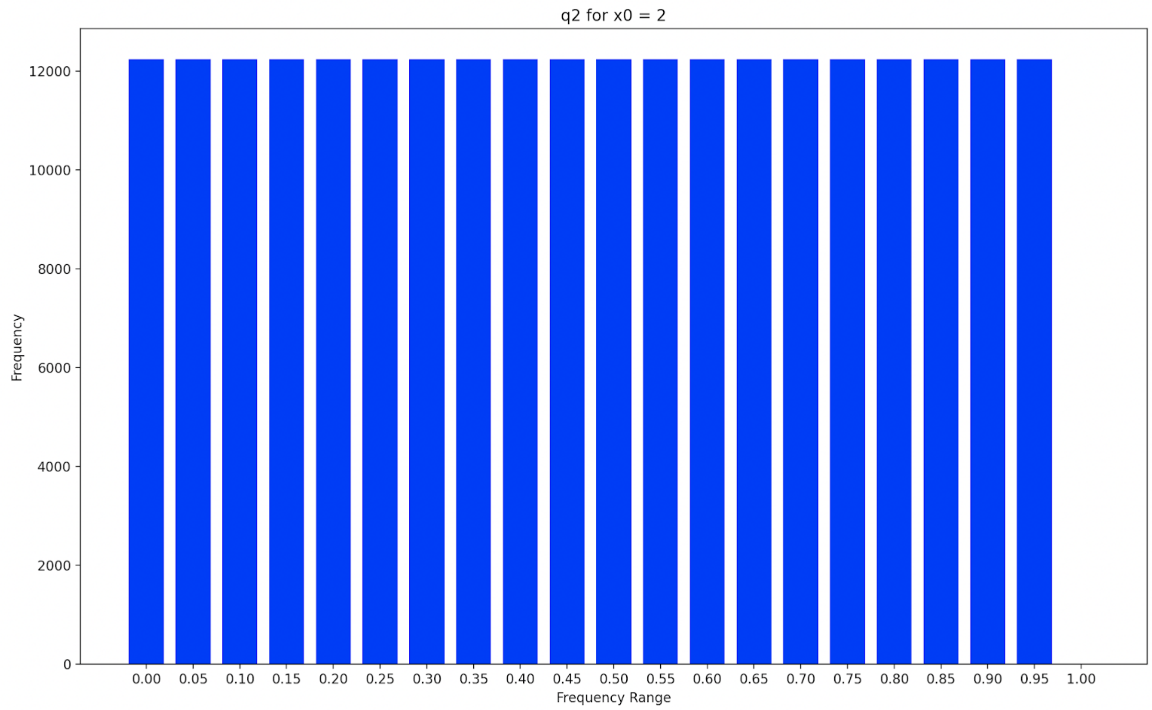
Problem II:
(* 0.00 indicates a range between 0.00 - 0.05)
$With\ a\ =\ 1597:$
$Tabulation:$

$Range$	$For\ x_0 = 1$	$For\ x_0 = 2$	$For\ x_0 = 3$	$For\ x_0 = 4$	$For\ x_0 = 5$
0.00 - 0.05	12249	12249	12249	12249	12249
0.05 - 0.10	12247	12247	12247	12247	12247
0.10 - 0.15	12247	12247	12247	12247	12247
0.15 - 0.20	12247	12247	12247	12247	12247
0.20 - 0.25	12247	12247	12247	12247	12247
0.25 - 0.30	12248	12248	12248	12248	12248
0.30 - 0.35	12247	12247	12247	12247	12247
0.35 - 0.40	12247	12247	12247	12247	12247
0.40 - 0.45	12247	12247	12247	12247	12247
0.45 - 0.50	12247	12247	12247	12247	12247
0.50 - 0.55	12248	12248	12248	12248	12248
0.55 - 0.60	12247	12247	12247	12247	12247
0.60 - 0.65	12247	12247	12247	12247	12247
0.65 - 0.70	12247	12247	12247	12247	12247
0.70 - 0.75	12247	12247	12247	12247	12247
0.75 - 0.80	12248	12248	12248	12248	12248
0.80 - 0.85	12247	12247	12247	12247	12247
0.85 - 0.90	12247	12247	12247	12247	12247
0.90 - 0.95	12247	12247	12247	12247	12247
0.95 - 1.00	12247	12247	12247	12247	12247
1.00 - 1.05	0	0	0	0	0

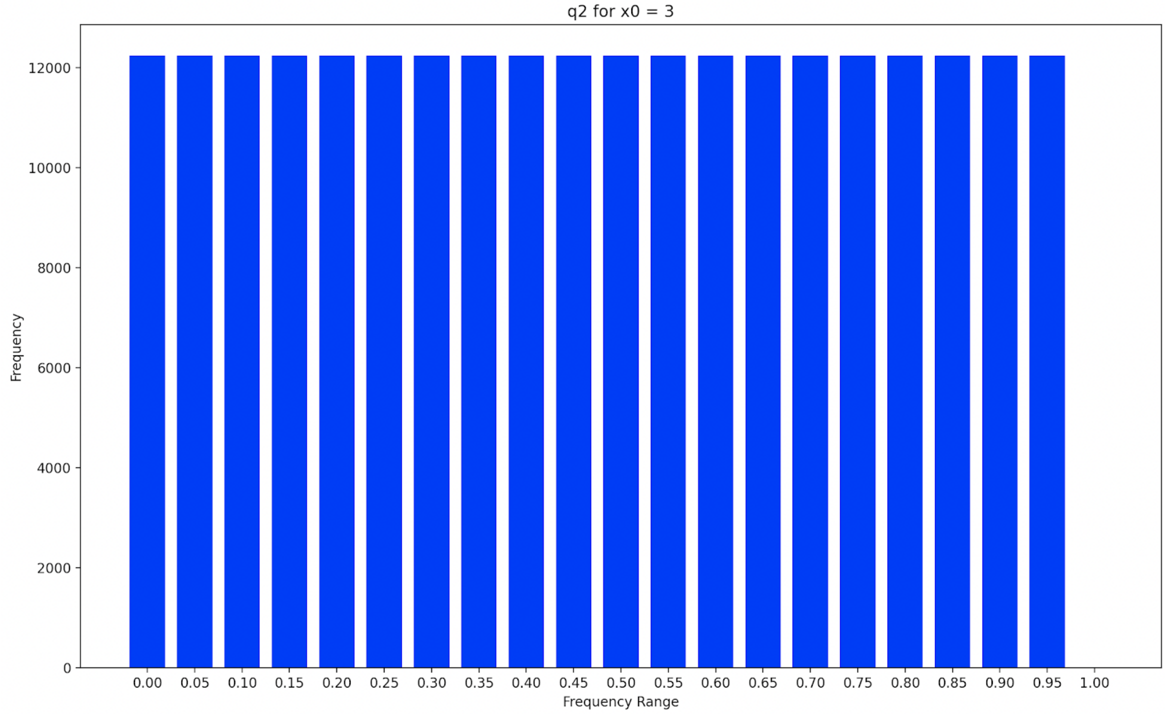
$For\ x_0 = 1:$



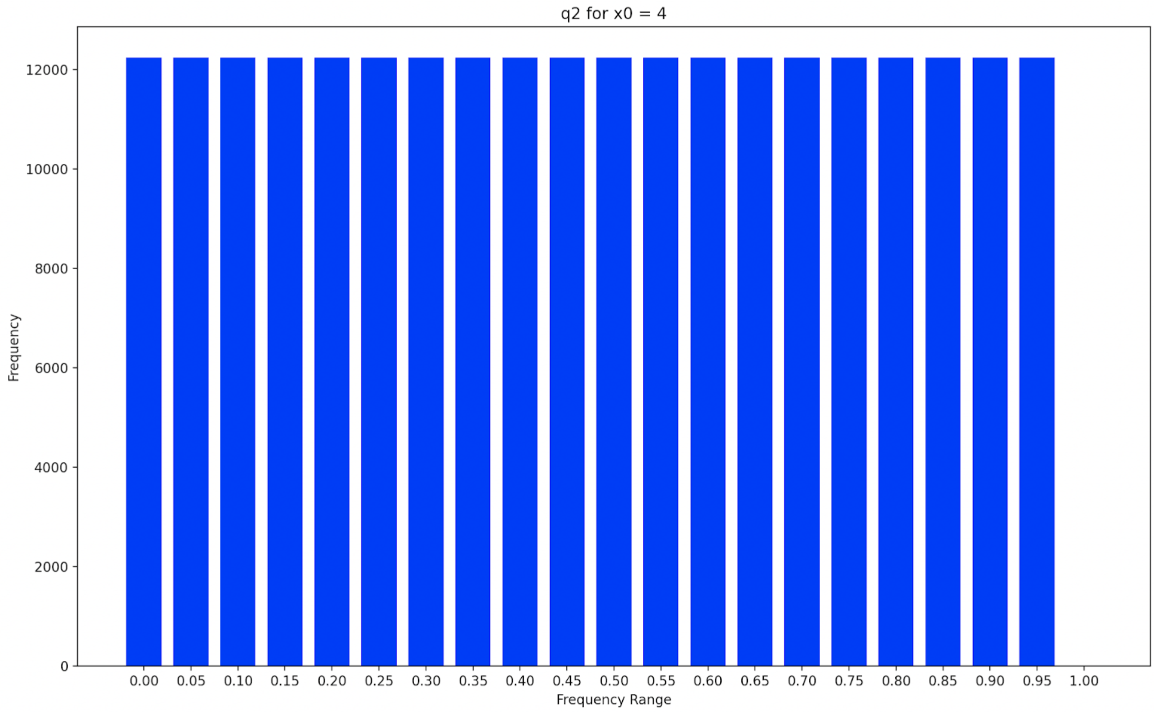
$For\ x_0 = 2:$



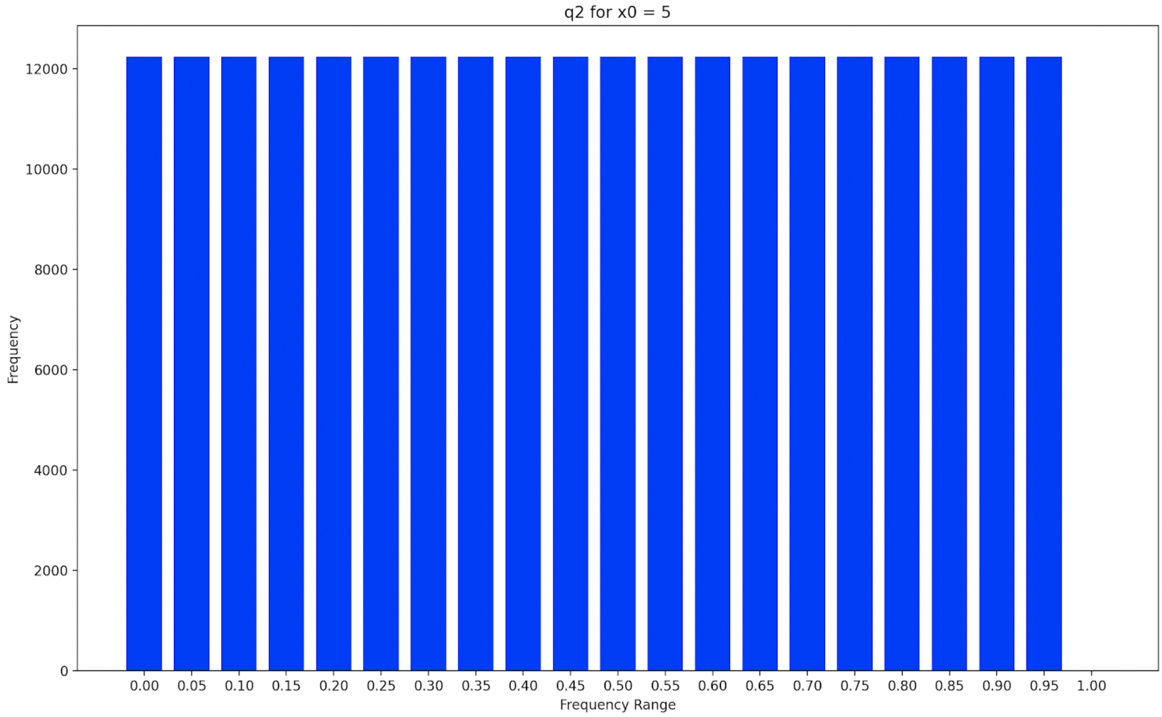
$For\ x_0 = 3:$



$For\ x_0 = 4:$



$For\ x_0 = 5:$

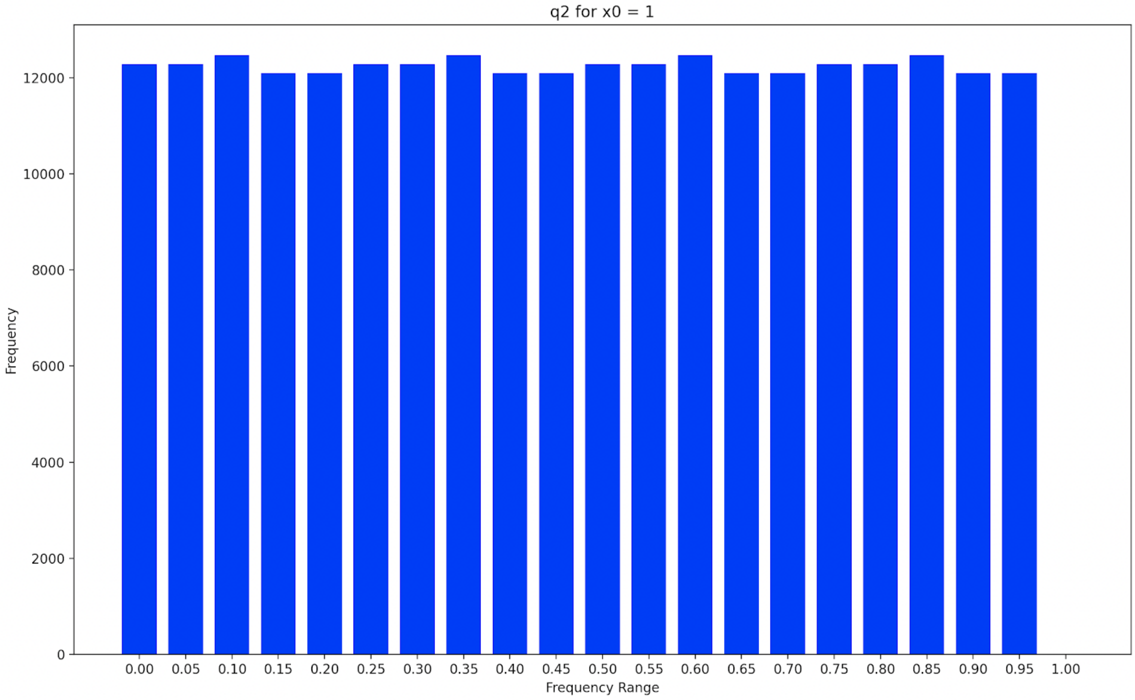


q2 for x0 = 5
Frequency
12000
10000
8000
6000
4000
2000
0
0.00 0.05 0.10 0.15 0.20 0.25 0.30 0.35 0.40 0.45 0.50 0.55 0.60 0.65 0.70 0.75 0.80 0.85 0.90 0.95 1.00
Frequency Range

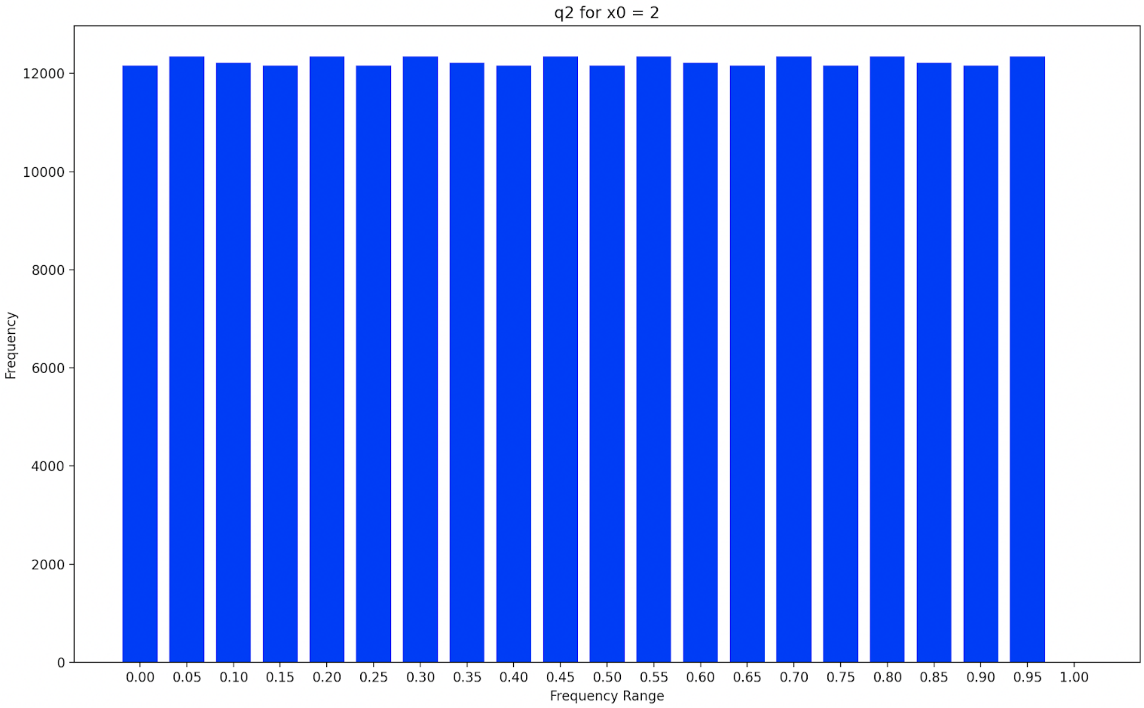
$With\ a\ =\ 51749:$
$Tabulation:$

$Range$	x_1	x_2	x_3	x_4	x_5
0.00 - 0.05	12286	12160	12286	13609	12223
0.05 - 0.10	12285	12348	12222	13608	12285
0.10 - 0.15	12474	12222	12285	10206	12285
0.15 - 0.20	12096	12159	12222	13608	12222
0.20 - 0.25	12096	12348	12222	10206	12222
0.25 - 0.30	12285	12159	12285	13608	12222
0.30 - 0.35	12285	12348	12222	13608	12285
0.35 - 0.40	12474	12222	12285	10206	12285
0.40 - 0.45	12096	12159	12222	13608	12222
0.45 - 0.50	12096	12348	12222	10206	12222
0.50 - 0.55	12285	12159	12285	13608	12222
0.55 - 0.60	12285	12348	12222	13608	12285
0.60 - 0.65	12474	12222	12285	10206	12285
0.65 - 0.70	12096	12159	12222	13608	12222
0.70 - 0.75	12096	12348	12222	10206	12222
0.75 - 0.80	12285	12159	12285	13608	12222
0.80 - 0.85	12285	12348	12222	13608	12285
0.85 - 0.90	12474	12222	12285	10206	12285
0.90 - 0.95	12096	12159	12222	13608	12222
0.95 - 1.00	12096	12348	12222	10206	12222
1.00 - 0.05	0	0	0	0	0

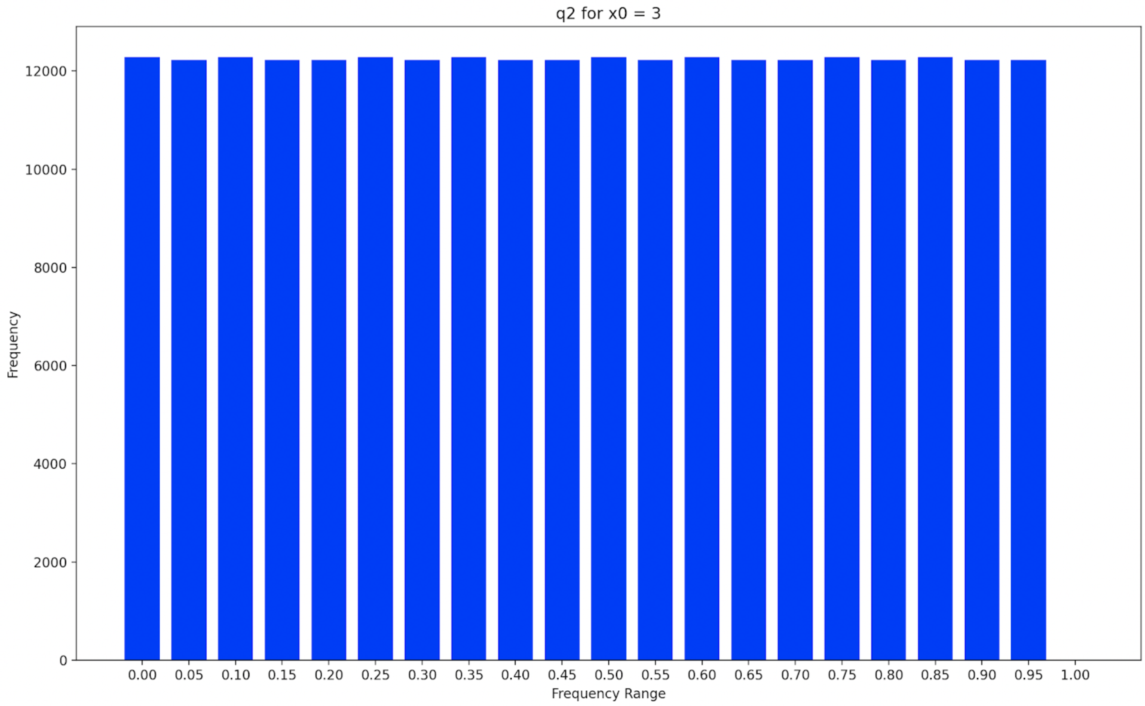
$For\ x_0 = 1:$



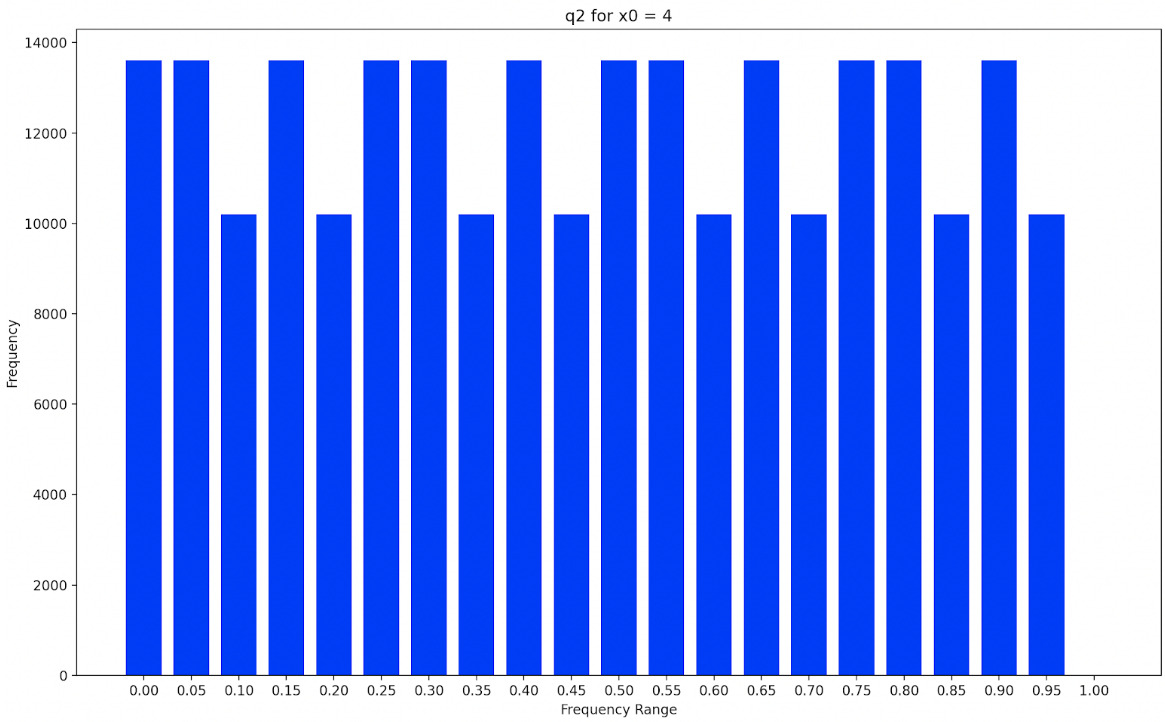
$For\ x_0 = 2:$



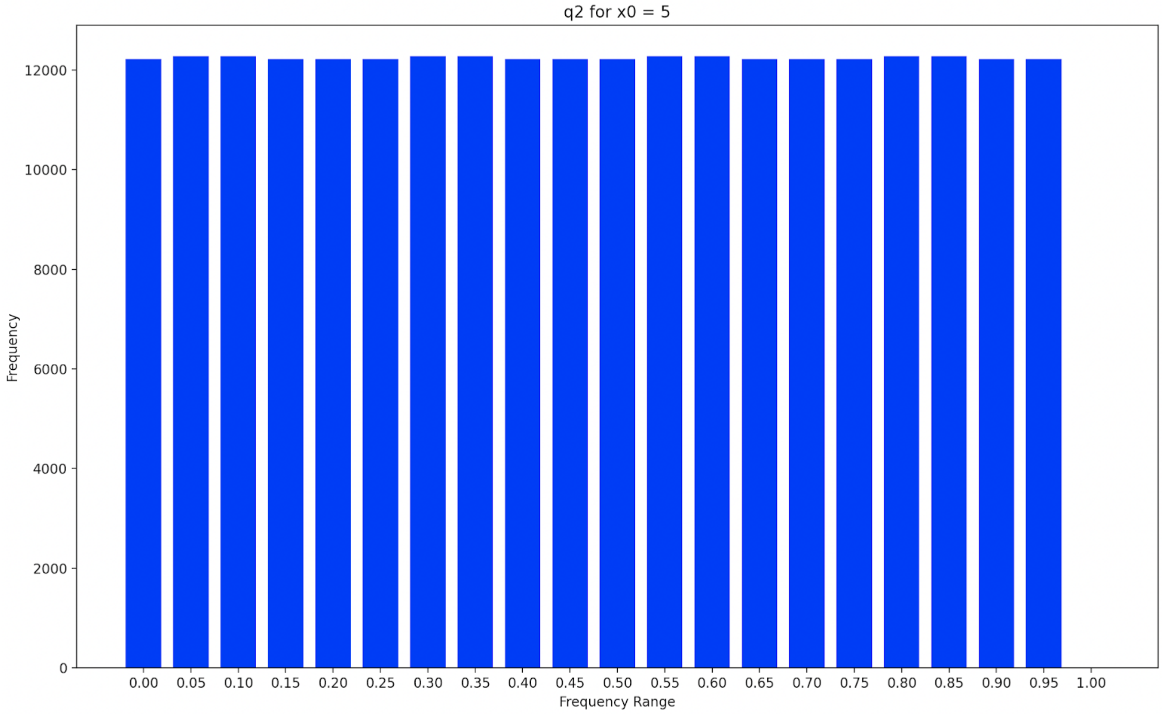
$For\ x_0 = 3:$



$For\ x_0 = 4:$



$For\ x_0 = 5:$

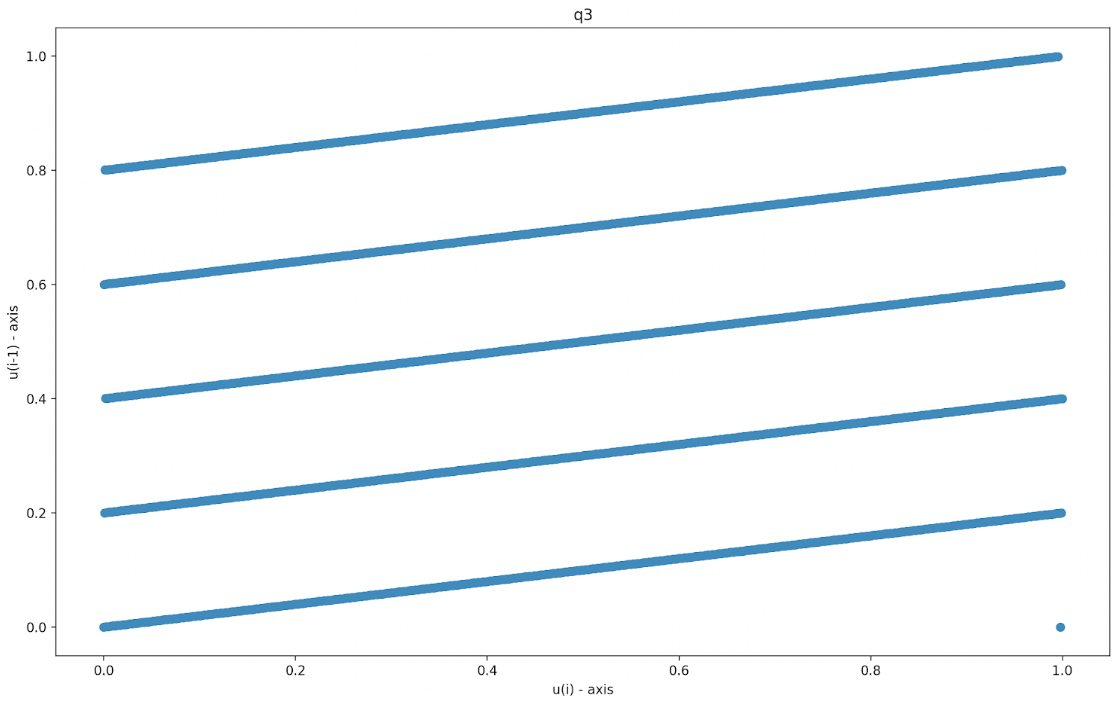


Observations:

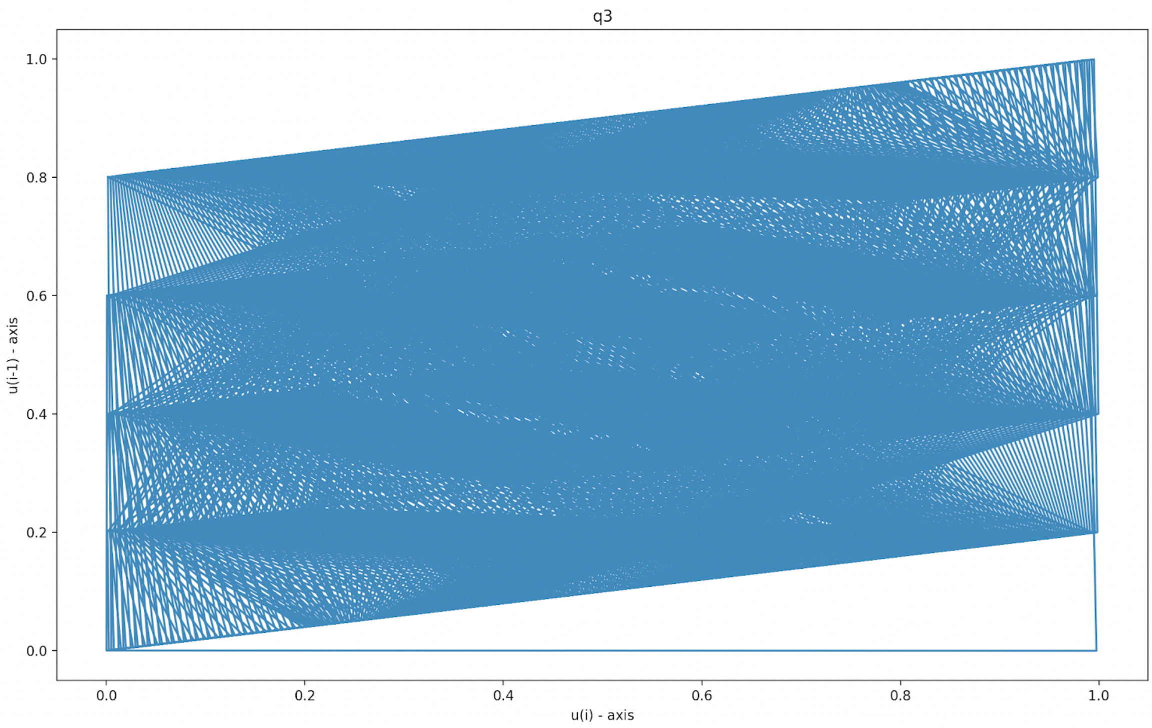
1. From the tabular data as well as the bar graphs, we can observe that the the distribution of the data appears to be independent of the chosen values of x_0 .
2. The frequency of various ranges is uniform, each section in the bar graph is more or less of the same height. *The exception being for $x_0 = 4\ and\ a\ =\ 51749$, two distinct heights for the bars can be observed.

Problem III:

$For\ x_0 = 3,\ a = 1229,\ b = 1,\ m = 2048,\ the\ plots\ obtained\ were\ as\ follows:$



I have used matplotlib.pyplot.scatter() function here.



I have used matplotlib.pyplot.plot() function here.