Exercise 2

In [171]:

```
h = open('sampleset.txt', 'r')

#Reading from the file
content = h.readlines()

#create a 100 size empty list
value = [None] * 100

for lines in content:
    for n in lines:
        if n.isdigit() == True:
            value.append(int(n))

sample = filter(None.__ne__, value)
samples = list(sample)
```

In [172]:

```
#using the help of https://stackoverflow.com/questions/16325988/factorial-of-a-large-number

def range_prod(low, high):
    if low + 1 < high:
        mid_number = (high + low) // 2
        return range_prod(low, mid)number * range_prod(mid_number + 1, high)
    if low == high:
        return low
    return low * high

def factorial(n):
    if n < 2:
        return 1
    return range_prod(1,n)
```

In [173]:

```
x = samples.count(1)
y = samples.count(0)
n = x + y
```

In [174]:

```python
#using the help of "https://medium.com/@tharakau/what-is-binomial-distribution-and-how-to-p

def bernoulliProb(n,x,p):
    prob = p**x * (1-p)**(n-x)
    return prob

Prob_value_list = []

theta = np.arange(0.0, 1.01, 0.01)
#print(theta)
large_value = -9999

for i in theta:
    Prob_value = bernoulliProb(n, x, i)
    #print("theta value: " , i , "Probability" , Prob_value , "\n")
    Prob_value_list.append(Prob_value)
    if Prob_value > large_value:
        large_value = Prob_value
        final_theta = i

print("Theta : ", final_theta , "Likelihood/probability : ", large_value)
```

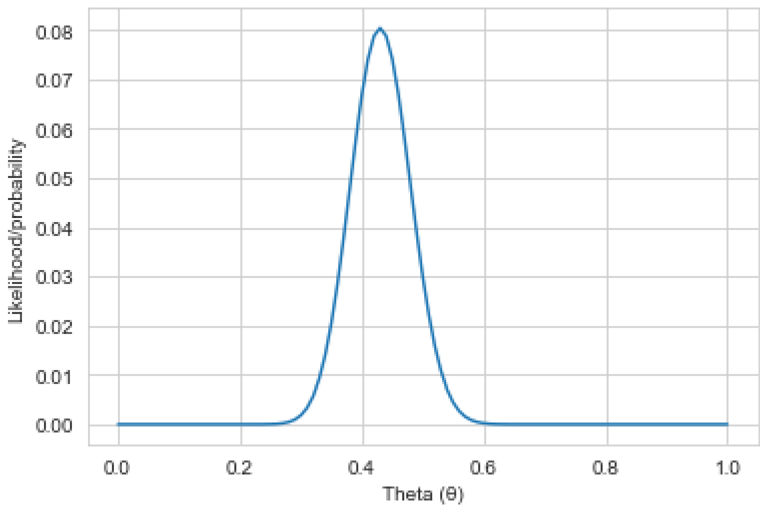
```
Theta :  0.43 Likelihood/probability :  2.1086784671978072e-30
```

In [178]:

```python
import matplotlib.pyplot as plt

plt.plot(theta, Prob_value_list)
plt.xlabel('Theta (θ)')
plt.ylabel('Likelihood/probability')
plt.show()

## In the graph, most of the theta value is between approximately 0.3 to 0.55. From 0.33, t
## and it reaches the highest peak when the theta value is 0.43. After that, it decreases a
```



In [176]:

```python
from itertools import permutations
from math import comb

def binomialProb(n,x,p):
    n_factorial = factorial(n)
    x_factorial = factorial(x)
    n_minus_x_factorial = factorial(n-x)
    result = n_factorial / (x_factorial * n_minus_x_factorial)
    prob = result * p**x * (1-p)**(n-x)
    return prob

Prob_value_list = []

theta = np.arange(0.0, 1.01, 0.01)
#print(theta)
large_value = -9999

for i in theta:
    #Prob_value = np.random.binomial(n, i)
    Prob_value = binomialProb(n, x, i)
    #print("theta value: " , i , "Probability" , Prob_value , "\n")
    Prob_value_list.append(Prob_value)
    if Prob_value > large_value:
        large_value = Prob_value
        final_theta = i

print("Theta : ", final_theta , "Likelihood/probability : ", large_value)
```

```
Theta :  0.43 Likelihood/probability :  0.08037551216200409
```

In [182]:

```python
import matplotlib.pyplot as plt

plt.plot(theta, Prob_value_list)
plt.xlabel('Theta (θ)')
plt.ylabel('Likelihood/probability')
plt.show()

## In the graph, most of the theta value is between approximately 0.3 to 0.55. From 0.33, t
## and it reaches the highest peak when the theta value is 0.43. After that, it decreases a
```