Given is a realization (*x1, …., xn*) of samples of n = 100 coin flips. The values xi = 1 represent heads and xi = 0 tails. We are interested in estimate the unknown parameter θ that is associated with the probability of heads. In order to approximate θ consider two different statistical models:

- it is assumed for the statistical model that the underlying $X_1, …., X_n$ are independent and identical distributed random variables following the Bernoulli distribution
- it is assumed for the statistical model that the underlying $X_1, …., X_n$ are independent and identical distributed random variables following the Binomial distribution.

Estimate θ using the Maximum Likelihood Method for the two statistical models. Compare the resulting values and comment on the difference.

Solution:

If a random sample *x1, ……, xn* of n observations is drawn from a Bernoulli distribution with parameter θ, this leads to the following likelihood and log-likelihood functions [1]:

$L(\theta) = (1-\theta)^{(1-x1)} * \theta^{x1} \ldots\ldots\ldots\ldots (1-\theta)^{(1-xn)} * \theta^{xn} = \prod_{i=1}^{n}(1-\theta)^{(1-xi)} * \theta^{xi}$ ……… (1)

Now, in the "sampleset.txt" the number of heads (i.e., 1) is *43*. So, we can write the equation (1) as follows:

$$L(\theta) = (1-\theta)^{(100-43)} * \theta^{43} \quad \ldots\ldots\ldots\ldots\ldots (2)$$

Take the log of (2), we got:

$\ln(L(\theta)) = \ln((1-\theta)^{57} * \theta^{43}) = 43 * \ln\theta + 57 * \ln(1-\theta)$ …………………… (3)

Take the derivative of the likelihood function equation (2) and setting it to 0, we found,

$$\frac{d}{d\theta}L(\theta) = \frac{d}{d\theta}(43 * \ln\theta + 57 * \ln(1-\theta))$$

$$= \frac{43}{\theta} - \frac{57}{1-\theta}$$

As $\frac{d}{d\theta}L(\theta) = 0$ so:

$$\frac{43}{\theta} - \frac{57}{1-\theta} = 0$$

- ⇨ $\frac{43*(1-\theta) - 57*\theta}{\theta*(1-\theta)} = 0$
- ⇨ $43 - 43*\theta - 57*\theta = 0$
- ⇨ $100*\theta = 43$
- ⇨ $\theta = \frac{43}{100}$
- ⇨ $\theta = 0.43$

Assume that X is a binomial distribution observation, $X \sim Bin\ (n, \theta)$, where n is known and θ is to be estimated. The probability function is as follows [2]:

$$L\ (x;\ \theta) = \frac{n!}{x! * (n-x)!} * \theta^{x} * (1-\theta)^{n-x}$$

Now [3],

$L\ (x1,\ ...,\ x100;\ \theta) = P_{X1\,...\,X100}\ (x1, \ldots, x100;\ \theta)$

$= P_{X1}(x1;\ \theta) \ldots . P_{x100}(x100;\ \theta)$

$= \binom{100}{1}^{43} * \binom{100}{0}^{57} * \theta^{43} * (1-\theta)^{57}$ [given that the number of head is 43 and tail is 57]

$= 1\ x\ 10^{86} * \theta^{43} * (1-\theta)^{57}$

We may take the derivative and set it to zero to discover the value of θ that maximizes the likelihood function. We have got:

$\frac{d}{d\theta} L(x1,\ ...,\ x100;\ \theta) = 1\ x\ 10^{86} * (43 * \theta^{42} * (1-\theta)^{57} - 57 * \theta^{43} * (1-\theta)^{56}) = 0$

- ⇨ $43 * \theta^{42} * (1-\theta)^{57} - 57 * \theta^{43} * (1-\theta)^{56} = 0$
- ⇨ $\frac{43}{57} * \frac{(1-\theta)^{57}}{(1-\theta)^{56}} = \frac{\theta^{43}}{\theta^{42}}$
- ⇨ $\frac{43}{57} * (1 - \theta) = \theta$
- ⇨ $\frac{43}{57} - \frac{43 * \theta}{57} = \theta$
- ⇨ $\frac{43}{57} = \theta + \frac{43 * \theta}{57}$
- ⇨ $43 = 57 * \theta + 43 * \theta$
- ⇨ $43 = 100 * \theta$
- ⇨ $\theta = \frac{43}{100} = 0.43$

The difference between θ in Bernoulli distribution and Binomial distribution are *0*. The binomial is, after all, the outcome of *n* separate Bernoulli trials.

The Bernoulli distribution, when $n = 1$, is a variant of the binomial distribution. X ~B (*1*, *p*) is equivalent to X ~Bernoulli in terms of symbolism (*p*). Any binomial distribution, B (*n*, *p*), is the sum of n separate Bernoulli trials, Bernoulli(*p*), all of which have the same probability *p* [4].

References:

[1]. Maximum likelihood method (ML), https://www.uni-kassel.de/fb07/index.php?eID=dumpFile&t=f&f=2722&token=79679e59f57ec8195642cdfd6ad1ea6327df6f78

[2]. Maximum-likelihood (ML) Estimation, https://online.stat.psu.edu/stat504/lesson/1/1.5

[3]. Maximum Likelihood Estimation, Example 8.8, https://www.probabilitycourse.com/chapter8/8_2_3_max_likelihood_estimation.php

[4]. Binomial distribution, https://en.wikipedia.org/wiki/Binomial_distribution

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]:

```
#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)
```

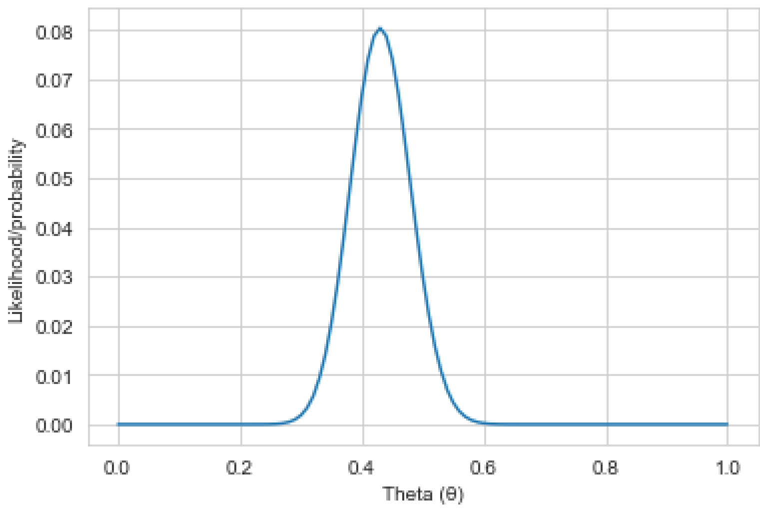
```
Theta :  0.43 Likelihood/probability :  2.1086784671978072e-30
```

In [178]:

```
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]:

```
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]:

```
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
```