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 X1, …., Xn are independent and identical distributed random variables following the Bernoulli distribution
* it is assumed for the statistical model that the underlying X1, …., Xn 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 (θ) = (1 – θ) (1 – x1) \* θx1 ………… (1 – θ) (1 – xn) \* θxn =* ...……. (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 (θ) =*  ……………….……… (2)

Take the log of (2), we got:

ln (L (θ)) = ln = 43 \* ln + 57 \* ln (1 – ) ………………….… (3)

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

*=*

As so:

*= 0*

* *= 0*

Assume that is a binomial distribution observation, , where is known and is to be estimated. The probability function is as follows [2]:

*L (x; ) =*

Now [3],

*L (x1, …, x100; =*

*=*

*=*  [given that the number of head is 43 and tail is 57]

= *1 x \**

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

*L(x1, …, x100; = 1 x \* = 0*

* *= 0*
* *\* =*
* *\* (1 - =*
* *- =*
* *= +*
* *43 = 57 \* + 43 \**
* *43 = 100 \**
* *= = 0.43*

The difference between 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>