**Background of the Study**

Parameter estimation is defined as a process used to make the inference of population parameters based on information from a sample taken from the entire population. Parameters are descriptive measures that represent the entire population. Since those parameters are unknown, it becomes difficult to measure the entire population.

A good estimator must have the following properties:

• Efficient: An estimator is said to be efficient if it has the lowest mean squared error

compared to others.

• Unbiased: An estimator is unbiased if it neither underestimates nor overestimates the

parameter. This means its mean should be equal to the true parameter.

• Consistency: An estimator is said to be consistent if as the sample size increases, the

estimator converges to the parameter it estimates.

• Minimum variance: A good estimator is one with the least variance compared to others

a) **MoM and MLE Estimation**

In descriptive statistic, we learned sample mean, sample variance etc. By calculating those values, in fact, we try to understand for example the location of the mean in the population by looking at the mean of the sample or we try to understand the variance in the population by looking the variance of the sample.
Although the it is very useful to use descriptive statistic, we want to estimate the type of the distribution that the population have. We can estimate it by using some parameter estimation's techniques.

Parameter Estimation is :

- It is more advanced and more informative statistical analysis.
- It is used to make desicions under uncertainty, develop optimal strategies, forecast, evaluate and control performance.

The following probability density function is given:

**f(x) = (2 * θ^2) / (x^3), θ <= x < ∞**
**f(x) = 0, elsewhere**

We'll use two most known estimation methods to estimate the parameter of a distribution:

**MoM:** Method of Moments,

**MLE:** Maximum Likelihood Estimation.

Since the population cannot be known, the sample collected from this population is used.

The sample set is: X = {0.3, 0.6, 0.8, 0.9}

**<u>For Method of Moments (MoM), the calculations are:</u>**

With a generic explanation,

the $k^{th}$ population moment:

$$\mu_k = E[x^k]$$

the $k^{th}$ sample moment:

$$m_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k$$

Our samples come from a family of distribution, we choose such a member of this family whose properties are close to sample properties

So we assume a family.

$$\begin{aligned} \mu_1 &= m_1 \\ &\vdots \\ \mu_k &= m_k \end{aligned}$$

Then we solve this system.

For our case: we begin by the first moment.

As we know,

$$M_1 = m_1$$

$$M_1 = E[X] = \int_{\theta}^{\infty} x f(x)\,dx = \int_{\theta}^{\infty} 2\theta^2 x^{-2}\,dx = 2\theta^2 \left[\frac{x^{-1}}{-1}\right]_{\theta}^{\infty} = -2\theta^2 \left[x^{-1}\right]_{\theta}^{\infty}$$

Then, we reach:

$$= -2\theta^2\left[0 - \frac{1}{\theta}\right] = 2\theta$$

Since the first moment ($m_1$) is equal to the sample mean:

$$m_1 = \bar{X}, \quad 2\hat{\theta} = \bar{X}$$

$$\hat{\theta} = \frac{\bar{X}}{2}$$

Then by finding sample mean, we can reach the extimated theta with this formula.

Let find the sample mean:

$$\bar{X} = \frac{0.3 + 0.6 + 0.8 + 0.9}{4} = 0.65$$

We reach our estimated parameter,

$$\hat{\theta} = 0.325$$

Hence, the **estimated parameter theta is 0.325 by using MoM.**

**<u>For Maximum Likelihood Estimation (MLE), the calculations are:</u>**

Given n random variables X1, X2, ..., Xn with the corresponding sample values x1, x2, ..., xn with a probability density function f(X1, ..., Xn, ω) where ω is a vector of parameters, the likelihood function of that sample is defined as

$$L(\omega | x1, x2, ..., xn) = f(x1, x2, ..., xn | \omega) = L(\omega).$$

Then, as we've known:

$$f(x_1, x_2, \dots, x_n) = \prod_{i=1}^{4} 2\theta^2 x_i^{-3} \quad \text{for} \quad \theta \leq x_i < \infty$$

First of all, X: $x_1, x_2, \dots, x_n$ , $f(X) = f(x_1, x_2, \dots, x_n)$.

Then for convenience of the calculation, we will take "ln" of both sides.

With the logarithmic property, we can say ln(a*b) = ln(a) + ln(b), so when we take "ln" of both sides the multiplication can be written as a summation:

Since $X: x_1, x_2, \dots, x_n$

$$f(X) = \prod_{i=1}^{4} 2\theta^2 x_i^{-3} \implies \ln f(X) = \ln\left(\prod_{i=1}^{4} 2\theta^2 x_i^{-3}\right) = \sum_{i=1}^{4} \ln(2\theta^2 x_i^{-3})$$

$$= \sum_{i=1}^{4} \ln 2 + \ln \theta^2 - 3\ln x_i = 4\ln 2 + 8\ln\theta - 3\sum_{i=1}^{4} \ln x_i$$

Our purpose is to estimate the maximum likelihood, so we need to consider the partial derivation of f(x) with respect to parameter theta. Since the parameter does not change we can take derivative of ln(f(x)) directly. To find the maximum, we need to find the estimated theta when the derivation is zero (if critical points exist).

$$\frac{\partial \ln f(x)}{\partial \theta} = \frac{8}{\theta} \neq 0$$

As we see, it cannot be 0. But when it is undefined, where theta = 0, there can be critical point. However, in order to be a critical point, this point must be defined in the original function. Since theta cannot be 0 for original function (because of ln(theta)), there is no critical point.

Then lets turn back f(x):

$$f(X) = \prod_{i=1}^{4} 2\theta^2 x_i^{-3} \quad \text{for} \quad \theta \le x_i < \infty$$

It is clearly seen that f(x) is maximum when theta has its maximum value. Since theta <= $x_i$ for i = 1, 2, 3, 4; the maximum value theta can have is **0.3**.

So the **estimated theta is 0.3 by using MLE.**

## b) <u>Population Generation</u>

Let apply the inverse transformation method to generate random samples of the given distribution.

Procedure to generate X:

1) Obtain a standard uniform random variable (U) from a random number generator.
2) Compute X = $F^{-1}$(U)

Let look the calculation:

First of all, in order to apply inverse transformation, we need to find the cumulative density function of f(x) which denoted as F(x).

$$F(x) = \int_{-\infty}^{x} \frac{2\theta^2}{a^3}\, da = \int_{\theta}^{x} \frac{2\theta^2}{a^3}\, da = 2\theta^2 \int_{\theta}^{x} a^{-3}\, da = 2\theta^2 \left[ \frac{a^{-2}}{-2} \right]_{\theta}^{x}$$

$$F(x) = -\theta^2 \left[ a^{-2} \right]_{\theta}^{x} = -\theta^2 \left[ x^{-2} - \theta^{-2} \right] = 1 - \frac{\theta^2}{x^2}$$

Then we can apply the inverse transformation method.

$$X = F^{-1}(U)$$

Then we can solve the equation for X. As a result we reach a random variable whose cumulative distribution function is given.

$$F(x) = 1 - \frac{\theta^2}{x^2} = U$$

$$1 - U = \frac{\theta^2}{x^2} \Rightarrow x^2 = \frac{\theta^2}{1-U}$$

$$X = \frac{\theta}{\sqrt{1-U}}$$

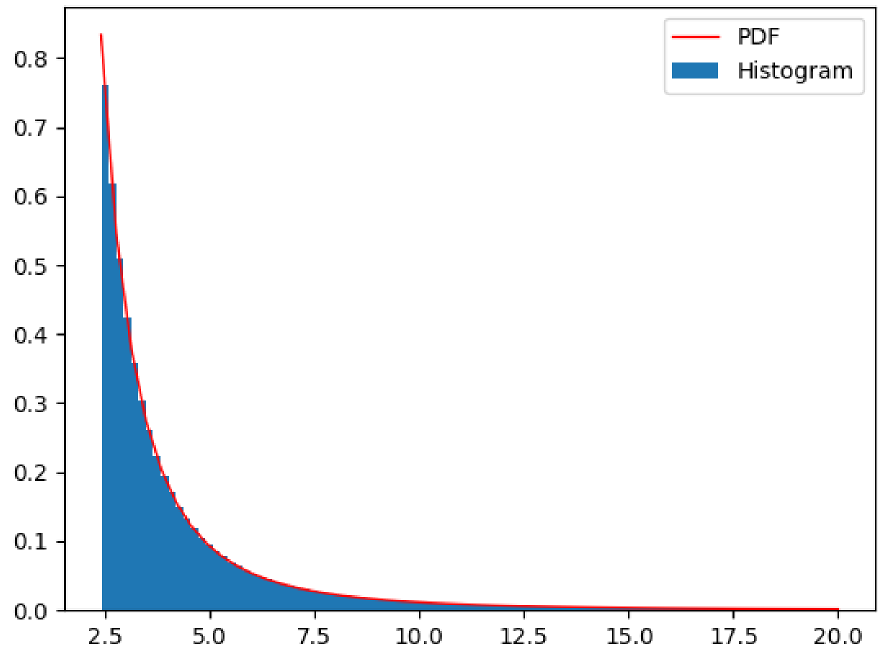
So by using this formula, now we can generate random variables X which has the given CDF.

## c) <u>Experiment Simulation</u>

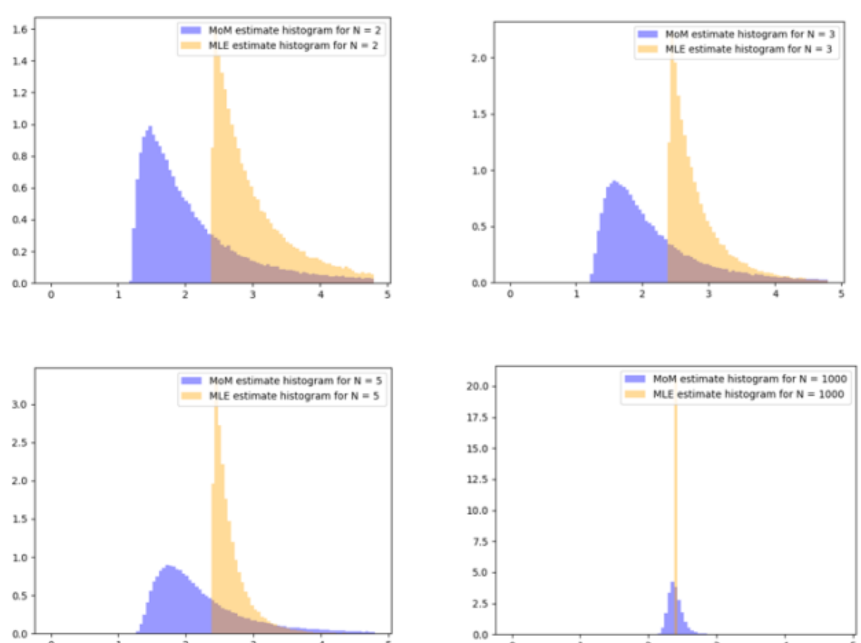
In Python, I created a population simulation with 10M of random variables. Then we create a sample size array which holds the size of the following samples that would be choosen randomly from the simulated population.

To see the effect of size of the samples, the array elements are in the increasing order. Our aim is to see the effects of the methods of moments and most likelihood estimation with the varying sample size.

The PDF of the distribution:




Some graphs from the experiment:



Çağın Tunç
ID: 280201090

**Results**:

We simulate data and estimate the theta parameters using Python software by setting the initial theta = 2.4. The simulation study is carried out to estimate the theta parameter. As the sample size increases, both of the methods estimates the real theta value which is 2.4. Hence, sampling error is decreasing while sample size increase. These two methods are compared based on the standard deviation of the estimates. From the obtained results, the maximum likelihood method is found to be the best method because of its minimum standard deviation. It also gives more efficient and less biased estimates compared to those given by moment method. We can see method of moment theta estimation spread its data throughout the graphs, it has more outliers than it was in the maximum likelihood estimation has. So we can understand MLE is better than MoM by just using simple descriptive statistics. We conclude that maximum likelihood method is better than the method of moment. Therefore, I would prefer to use the Maximum Likelihood Estimation.