# STATISTICAL METHODS FOR DATA SCIENCE CS 6313-001 FALL 2019

**Mini Project #3**

**Participant : Reetish Chand Guntakal Patil (RXG190006)**

## Question 1.

## 

## 

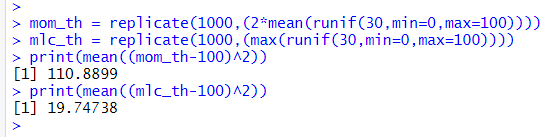
1. Computing the mean squared error :

We compute the mean squared error by finding difference between the estimates of random sample and the estimator whole squared.

MSE= E[(Obs-exp)^2]

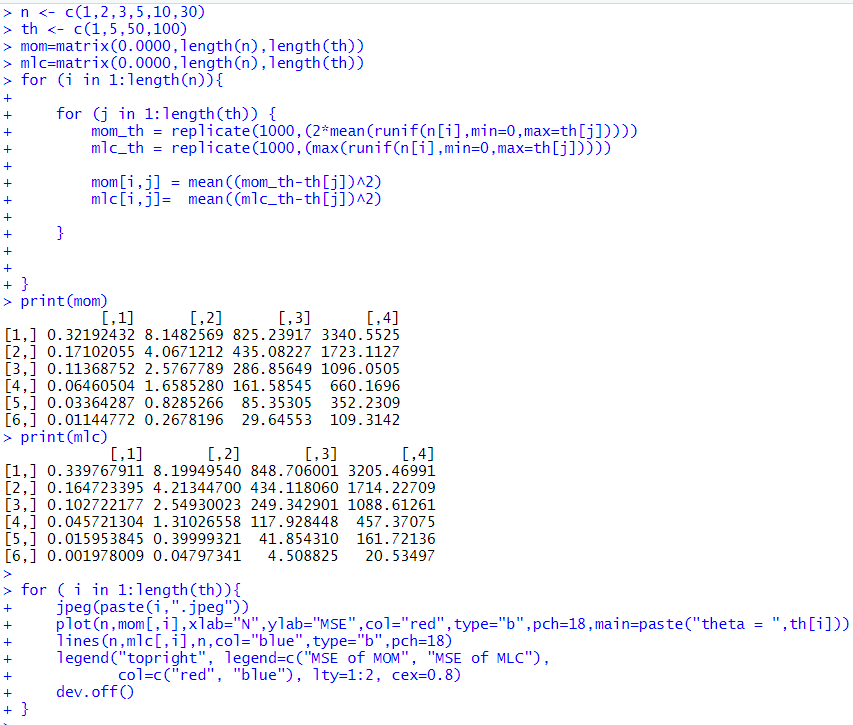
In R we use mse function of hyroGOF package

1. To compute ϴ1 and ϴ2 : Let n=30 and ϴ=100.



Observation:

The mean squared error for maximum likelihood estimator is less than that of method of moments.

1. Repeating the above for all given n, ϴ combinations :

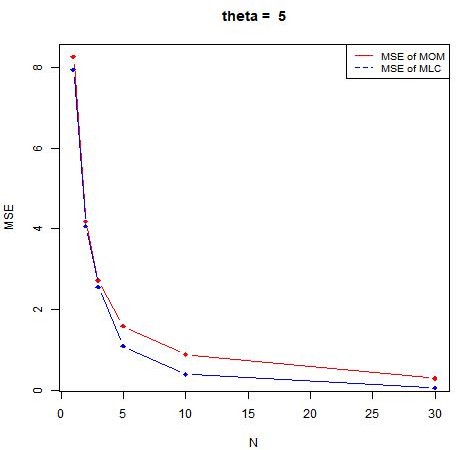
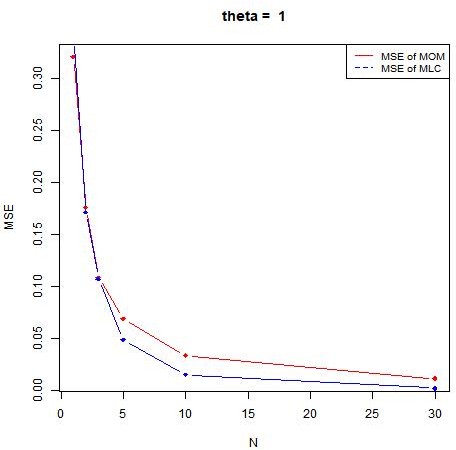
We write a nested for loop which calls the above function with all possible n, ϴ

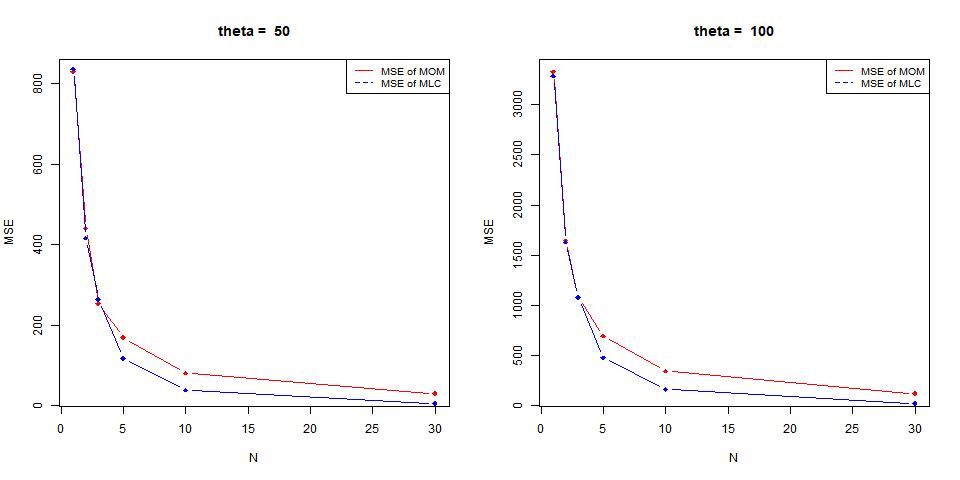
combinations. A simple plot() is used to plot the graph. We use Jpeg() to create a

.jpeg image file in our working directory containing each plot of mean squared error of method of moments and maximum likelihood estimator.

For each theta value N values are plotted on the x axis and mean square errors are plotted on the y axis.

We get the following :





Observations:

From all the line graphs it can be observed that the mean squared error for maximum likelihood estimator is less compared to that of method of moments.

1. according to the results for all the combinations of n and ϴ, the mean squared error for Maximum likelihood estimator is less than the Method of moments estimation thus, indicating that the Maximum likelihood is a better estimate than method of moment.

The estimation is dependent on n and ϴ because both Maximum likelihood and Method of moments have different estimates for ϴ. As the sample size n changes the result also changes accordingly.

## Question 2:

## 

* 1. ∑Given f(x) = ϴ/x ϴ+1

Applying log we get log( ϴ/x ϴ+1)

= ∑1 to n (log( ϴ/x ϴ+1))

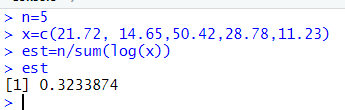
= ∑1 to n (log( ϴ)-log(x ϴ+1))

= ∑1 to n (log( ϴ)- (ϴ +1)log(x))

=n log ϴ - (ϴ+1) ∑1ton log(x) Differentiating wrt to ϴ (n/ ϴ) - ∑1 to n log(x) = 0

Therefore ϴ = n/ ∑1 to n log(x)

* 1. Solving ϴ:



As we can see the estimated value of ϴ is 0.3233874

* 1. Maximizing the log likelihood function:



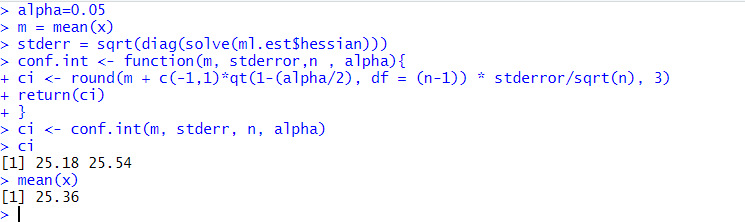
Observation:

The par value and the estimated ϴ value from (b) are the same.

The hessian value is 47.81158

* 1. to approximate standard error of the maximum likelihood estimate and an approximate 95%

confidence interval for ϴ:



As we can see the mean of given x values lies in the computed confidence level. Therefore, we can say that given sample is a good approximator for the given f(x).