A Simulation Study on Modified Minimum Distance Estimators

A project work completed as a part of the course MTH511A

Submitted by

Anis Pakrashi (211264)

Krishnendu Paul (211322)

Rahul Ghosh Dastidar (211353)

Souraj Mazumdar (211393)

Under the supervision of

Dr. Arnab Hazra

Department of Mathematics and Statistics

Indian Institute of Technology, Kanpur

Acknowledgement

We present our project report on **"A Simulation Study on Modified Minimum Distance Estimators"**. However, as mere students of statistics, any achievement, in terms of a substantial project, would have been an uphill task, considering just our own efforts. It is apparent that there has always been a constant encouragement of noble minds for our efforts to bring this project to a successful completion. The satisfaction that accompanies the effectiveness of any task would be incomplete without mentioning those who made it possible for us to complete our desired project.

First and foremost, we want to express our sincere gratitude to our instructor **Dr. Arnab Hazra** for his constant help, support and advice throughout the project preparation, which acted as a catalyst for effective completion of our work. Without his valuable guidance and motivation, it would have been nearly impossible to work as a team and imbibe the practical aspects of the course **"MTH511A: Statistical Simulation and Data Analysis"**. Besides, we are thankful to all faculty members of our department and our seniors, because without their support at various stages, this project would not have materialized.

We are also thankful to our friends for their critical appreciation and their questions and counter-questions solved a huge spectrum of doubts related to our work. Last but not the least, we are grateful to our parents for their constant motivation on the way of materializing the project.

Anis Pakrashi

Krishnendu Paul

Rahul Ghosh Dastidar

Souraj Mazumdar

Abstract

In our project we review the paper *Modified minimum distance estimators: definition, properties and applications* by Talha Arslan, Sukru Acitas, Birdal Senoglu. Estimating the location and scale parameters of a distribution is one of the most crucial issues in Statistics.There exists various estimators such as maximum likelihood, method of moments and minimum distance (e.g.Cramér von Mises—CvM and Anderson Darling—AD), etc. However, in most cases, we cannot obtain closed forms of estimators because the estimating equations involve non-linear functions. Now, the numerical methods, used to obtain the estimates, may have some drawbacks such as multiple roots, wrong convergency, and non-convergency of iterations. In this paper, the authors adopted the idea of Tiku (Biometrika 54:155–165, 1967) into the CvM and AD methodologies with the intent of eliminating the aforementioned difficulties and obtaining closed form estimators of the parameters μ and σ. Resulting estimators are called as modified CvM (MCvM) and modified AD (MAD), respectively. Proposed estimators are expressed as functions of sample observations and thus their calculations are straightforward. We have performed Monte-Carlo simulation study to confirm their findings. We compared the efficiencies of the CvM and AD estimators with their modified counterparts, i.e. the MCvM and MAD, as well as the maximum likelihood and the modified least squares for the normal, extreme value and Weibull distributions for an illustration. A real data set has been used to show the implementation of the proposed estimation methodologies and hence obtain the corresponding estimates.

Keywords: Modified Minimum Distance, Cramer von Mises, Anderson Darling, Monte Carlo simulation

Contents

1 Introduction:

Minimum distance estimation methods are widely used by the practitioners from several fields. These estimation techniques are preferred when outliers are present in data. The fundamental approach in minimum distance estimation methods is to minimize the distance between the empirical cumulative distribution function (cdf) and theoretical (or population) cdf.

While talking about minimizing the distance between the empirical cdf and theoretical (or population) cdf we first require to define what is a "Statistical Distance". There are various statistical distances those are used in minimum distance estimation. Kolmogorov-Smirnov (K-S), Cramér-von Mises (CvM) and Anderson Darling (AD) are some well-known and widely used statistical distances which are used in both minimum distance estimation and nonparametric hypothesis testing. While evaluating these minimum distance estimators, most of the time the estimating equations involve nonlinear functions of the parameters. To solve these equations we then use numerical methods and various optimization techniques. But unfortunately many times these methods do not work as the methods do not converge, converge slowly or even converge to wrong root.

In the paper the authors suggest to linearize the nonlinear functions existing in the estimating equations around the expected values of the ordered observations.This approach helps to get rid of the aforesaid difficulties and produces modified versions of the CvM and AD estimators. In this project we are trying to implement those ideas through Simulation Studies and Real Data Analyses.The newness of this study is that the computations of the proposed estimators are straightforward which involve no iterative procedures and thus dramatically reduces the computational cost. It should be mentioned that this idea was adopted by the authors of this paper from Tiku (1967) in which the modified maximum likelihood (MML) methodology is proposed.

2 Estimation of the location and scale parameters:

In this section the notations will be introduced first. This will be followed by the description of minimum distance estimation methods CvM and AD estimators along with their modified versions.

2.1 Notations :

The notations used throughout this project are given as follows:
$F(\cdot)$: cumulative distribution function (cdf),
$f(\cdot)$: probability distribution function (pdf),
$f^{'}(\cdot)$: derivative of the f $(\cdot)$,
$x_{(i)}$: i -th ordered observation,
$z_{(i)}$: i -th standardized ordered observation, i.e., $z_{(i)} = \frac{x_{(i)}-\mu}{\sigma}$.
Here, our goal is to estimate the location parameter μ and scale parameter σ of a distribution.

2.2 The CvM and Modified CvM estimation :

Consider the objective function :

$$CvM = \sum_{i=1}^{n}[F(x_{(i)};\mu,\sigma) - \frac{2i-1}{2n}]^2 + \frac{1}{12n} \tag{1}$$

Minimizing the above objective function with respect to the parameters μ and σ , we can obtain the CvM estimators of μ and σ.

The estimating equations are as follows :-

$$\frac{\partial CvM}{\partial \mu} = -2\sum_{i=1}^{n} f(z_{(i)})[F(z_{(i)}) - \frac{2i-1}{2n}] = \sum_{i=1}^{n}(\frac{2i-1}{2n})f(z_{(i)}) - \sum_{i=1}^{n} f(z_{(i)})F(z_{(i)}) = 0 \tag{2}$$

and

$$\frac{\partial CvM}{\partial \sigma} = -2\sum_{i=1}^{n} z_{(i)}f(z_{(i)})[F(z_{(i)}) - \frac{2i-1}{2n}]$$

$$= \sum_{i=1}^{n} z_{(i)}(\frac{2i-1}{2n})f(z_{(i)}) - \frac{2}{\sigma}\sum_{i=1}^{n} z_{(i)}f(z_{(i)})F(z_{(i)}) = 0 \tag{3}$$

Since estimating equations (2) and (3) include nonlinear functions, estimates of μ and σ cannot be obtained explicitly. Therefore, we have to take help of numerical methods. However, as discussed in Introduction some computational difficulties may arise. Hence we proceed by following exactly the same steps as in MML methodology which is described below :-

Step 1: Standardized observations $z_i (= \frac{x_i-\mu}{\sigma})$ are ordered in ascending way, i.e. $z_{(1)} \leq z_{(2)} \leq z_{(3)} \leq .... \leq z_{(n)}$.

Step 2: The nonlinear functions present in equations (2) and (3) ; i.e.

$$h_1(z_{(i)}) = (\frac{2i-1}{2n})f(z_{(i)}) \text{ and } h_2(z_{(i)}) = f(z_{(i)})F(z_{(i)})$$

are linearized using the first two terms of Taylor series expansion around the expected values of the standardized order statistics $t_{(i)} = E(z_{(i)})$. So the linearized forms are −

$$h_1(z_{(i)}) \cong \alpha_{1i} + \beta_{1i}z_{(i)} \text{ and } h_2(z_{(i)}) \cong \alpha_{2i} + \beta_{2i}z_{(i)} \tag{4}$$

$$\alpha_{1i} = h_1(t_{(i)}) - t_{(i)}(\frac{2i-1}{2n})f^{'}(t_{(i)}), \quad \beta_{1i} = (\frac{2i-1}{2n})f^{'}(t_{(i)}),$$

$$\alpha_{2i} = h_2(t_{(i)}) - t_{(i)}[f^{'}(t_{(i)})F(t_{(i)}) + f(t_{(i)})^2] \text{ and}$$

$\beta_{2i} = [f^{'}(t_{(i)})F(t_{(i)}) + f(t_{(i)})^2]$.

Step 3: Using equations in (4) into the equations (2) and (3), we get modified versions of them as follows

$$\frac{\partial CvM^*}{\partial \mu} = \sum_{i=1}^{n}(\alpha_{1i} + \beta_{1i}z_{(i)}) - \sum_{i=1}^{n}(\alpha_{2i} + \beta_{2i}z_{(i)}) = 0 \qquad (5)$$

and

$$\frac{\partial CvM^*}{\partial \sigma} = \sum_{i=1}^{n} z_{(i)}(\alpha_{1i} + \beta_{1i}z_{(i)}) - \sum_{i=1}^{n} z_{(i)}(\alpha_{2i} + \beta_{2i}z_{(i)}) = 0 \qquad (6)$$

Step 4: The solutions of modified equations (5)and (6) are the following **MCvM** estimators

$$\hat{\mu}_{MCvM} = \overline{x}_w - \frac{\mathit{\Delta}}{\Lambda}\hat{\sigma}_{MCvM} \text{ and } \hat{\sigma}_{MCvM} = \frac{\sum_{i=1}^{n}\lambda_i(x_{(i)}-\overline{x}_w)^2}{\sum_{i=1}^{n}\delta_i(x_{(i)}-\overline{x}_w)} \qquad (7)$$

where,

$$\overline{x}_w = \frac{\sum_{i=1}^{n}\lambda_i x_{(i)}}{\Lambda}, \lambda_i = \beta_{2i} - \beta_{1i},\ \Lambda = \sum_{i=1}^{n}\lambda_i, \delta_i = \alpha_{1i} - \alpha_{2i} \text{ and } \mathit{\Delta} = \sum_{i=1}^{n}\delta_i \qquad (8)$$

2.3 The AD and Modified AD estimation

Consider the objective function :

$$AD = -n - n^{-1}\sum_{i=1}^{n}\{(2i-1)lnF(x_{(i)};\mu,\sigma) + (2n+1-2i)ln[1-F(x_{(i)};\mu,\sigma)]\} \qquad (9)$$

Minimizing the above objective function with respect to the parameters μ and σ , we can obtain the AD estimators of μ and σ.

The estimating equations are as follows :-

$$\frac{\partial AD}{\partial \mu} = \frac{1}{n}\sum_{i=1}^{n}(2i-1)\frac{f(z_{(i)})}{F(z_{(i)})} - \frac{1}{n}\sum_{i=1}^{n}(2n+1-2i)\frac{f(z_{(i)})}{1-F(z_{(i)})} = 0 \qquad (10)$$

$$\frac{\partial AD}{\partial \sigma} = \frac{1}{n}\sum_{i=1}^{n} z_{(i)}(2i-1)\frac{f(z_{(i)})}{F(z_{(i)})} - \frac{1}{n}\sum_{i=1}^{n} z_{(i)}(2n+1-2i)\frac{f(z_{(i)})}{1-F(z_{(i)})} = 0 \qquad (11)$$

Similar to the CvM estimation, the explicit solutions of equation (10) and (11) cannot be obtained. So we follow the same steps as in section 2.2 to obtain the MAD estimators of μ and σ.

Here the nonlinear functions are :-

$$h_1(z_{(i)}) = (2i-1)\frac{f(z_{(i)})}{F(z_{(i)})} \text{ and } h_2(z_{(i)}) = (2n+1-2i)\frac{f(z_{(i)})}{1-F(z_{(i)})}$$

Hence, $\alpha_{1i,}\beta_{1i,}\alpha_{2i}$ and $\beta_{2i,}$ in the linearized forms of the $h_1(z_{(i)})$ and $h_2(z_{(i)})$ are obtained as follows :-

$$\alpha_{1i} = h_1(t_{(i)}) - t_{(i)}[\frac{f^{'}(t_{(i)})F(t_{(i)})-f(t_{(i)})^2}{F(t_{(i)})^2}](2i-1)$$

$$\alpha_{2i} = h_2(t_{(i)}) - t_{(i)}[\frac{f^{'}(t_{(i)})(1-F(t_{(i)}))+f(t_{(i)})^2}{(1-F(t_{(i)}))^2}](2n+1-2i)$$

$$\beta_{1i} = [\frac{f^{'}(t_{(i)})F(t_{(i)})-f(t_{(i)})^2}{F(t_{(i)})^2}](2i-1)$$

$$\beta_{2i} = [\frac{f^{'}(t_{(i)})(1-F(t_{(i)}))+f(t_{(i)})^2}{(1-F(t_{(i)}))^2}](2n+1-2i)$$

Then, the MAD estimators of the location parameter μ and scale parameter σ are obtained as follows:

$$\hat{\mu}_{MAD} = \bar{x}_w - \frac{\varDelta}{\Lambda}\hat{\sigma}_{MAD} \text{ and } \hat{\sigma}_{MAD} = \frac{\sum_{i=1}^{n}\lambda_i(x_{(i)}-\bar{x}_w)^2}{\sum_{i=1}^{n}\delta_i(x_{(i)}-\bar{x}_w)} \qquad (12)$$

where, $\bar{x}_w$,Λ,$\varDelta$,$\lambda_i$and δ_i are defined as in equation (8).

3 Properties of the proposed estimators

Some properties of the modified estimators are as follows :-

- As $n \to \infty$,the differences $[h_1(z_{(i)})-(\alpha_{1i}+\beta_{1i}z_{(i)})] \to 0$ and $[h_2(z_{(i)})-(\alpha_{2i}+\beta_{2i}z_{(i)})] \to 0$. As a result, $\underset{n\to\infty}{lim}\frac{1}{n}\left|\frac{\partial MD}{\partial\mu} - \frac{\partial MD^*}{\partial\mu}\right| = 0$ and $\underset{n\to\infty}{lim}\frac{1}{n}\left|\frac{\partial MD}{\partial\sigma} - \frac{\partial MD^*}{\partial\sigma}\right| = 0$. Thus, the MCvM and MAD estimators of the location parameter μ and scale parameter σ are asymptotically equivalent to their counterparts CvM and AD estimators.

- If the underlying distribution is symmetric, then $t_{(i)} = -t_{(n-i+1)}$, $\delta_i = -\delta_{n-i+1}(i.e. \sum_{i=1}^{n}\delta_i = 0)$ and $\lambda_i = \lambda_{n-i+1}$. Hence, the MCvM or MAD estimator of the location parameter μ turns out to be -

$\hat{\mu} = \frac{\sum_{i=1}^{n}\lambda_i x_{(i)}}{\sum_{i=1}^{n}\lambda_i}$,which is clearly a weighted mean of the sample ordered observations with the weights $\lambda_i; i = 1, 2, ..., n$.

Note that small weights λ_i are assigned to the outlying observations, which makes the modified estimators robust to anomalies in the data.

- Another remarkable property of these methods is that structure of the closed forms of the resulting estimators remain same for any distribution of interest similar to the MML estimators.

4 Simulation Studies

Here we will perform Monte-Carlo Simulations to compare the efficiencies of the CvM and AD estimators, their modified counterparts (i.e. MCvM and MAD estimators), MLE (Maximum Likelihood Estimator) and MLSE (Modified Least Square Estimator). The MLSE of the parameters μ and σ are obtained as :-

$$\hat{\mu}_{MLS} = \bar{x}_w - \frac{\varDelta}{\Lambda}\hat{\sigma}_{MLS} \text{ and } \hat{\sigma}_{MLS} = \frac{\sum_{i=1}^{n}\lambda_i(x_{(i)}-\bar{x}_w)^2}{\sum_{i=1}^{n}\delta_i(x_{(i)}-\bar{x}_w)} \qquad (13)$$

where the steps of calculations for the $\hat{\mu}_{MLS}$, $\hat{\sigma}_{MLS}$ and the definitions of $\bar{x}_w, \Lambda, \varDelta, \lambda_i$ and δ_i are same as discussed in section 3. It should also be noted that in obtaining the MLS estimators, nonlinear functions $h_1(z_{(i)}) = (\frac{i}{n+1})f(z_{(i)})$ and $h_2(z_{(i)}) = f(z_{(i)})F(z_{(i)})$ exist in the corresponding estimating equations.

We will generate random samples from $N(0, 1)$, $EV(0, 1)$ and $Weibull(0, 1, 1.2)$ distributions. Clearly, here $\mu = 0$ and $\sigma = 1$ for all the distributions we will simulate from.

All the simulations will be conducted for 1000 Monte-Carlo runs . The sample sizes will be taken as n = 30 (small), n = 50 (moderate) and n = 100 (large) for each run.

4.1 Bias and efficiency comparisons

We will compare the CvM, AD, MCvM, MAD, ML and MLS estimators by calculating their Bias and Mean Squared Error (MSE) for each of the distributions and for the different sample sizes mentioned above. The Bias and MSE values are calculated using the following formulae :-

$$Bias(\hat{\theta}) = \theta - \frac{1}{N}\sum_{i=1}^{N}\hat{\theta}_i \text{ and } MSE(\hat{\theta}) = \frac{1}{N}\sum_{i=1}^{N}(\theta - \hat{\theta}_i)^2$$

where $N = 1000$, $\theta = \mu$ or σ and $\hat{\theta}_i$ stands for the CvM, AD, MCvM, MAD, ML or MLS estimates of the corresponding parameter at the i-th Monte-Carlo run.

We will also consider the deficiency (Def) criterion which is defined as the joint efficiency of the estimators μ and σ. Def values are calculated using the following formula :-

$$Def = MSE(\hat{\mu}) + MSE(\hat{\sigma})$$

4.1.1 Sampling from Normal(0,1) Distribution

We will now show the results of our simulation study for the different sample sizes :

- **Sample Size (n) = 30**

The following table shows the Bias and MSE of the different estimators of μ and σ and the corresponding Def values for our simulated sample –

```
[1] "Sample size =  30"
[1] "Observation matrix is "
     mu.bias mu.mse sigma.bias sigma.mse    def
MLE   0.0003 0.0348     0.0296    0.0188 0.0536
MLS  -0.0023 0.0380    -0.0775    0.0370 0.0751
CvM  -0.0020 0.0383     0.0251    0.0293 0.0677
MCvM -0.0022 0.0380    -0.0264    0.0291 0.0671
AD   -0.0007 0.0361    -0.0005    0.0219 0.0580
MAD  -0.0009 0.0361    -0.0177    0.0221 0.0582
```

$Table$: 1

Observations –

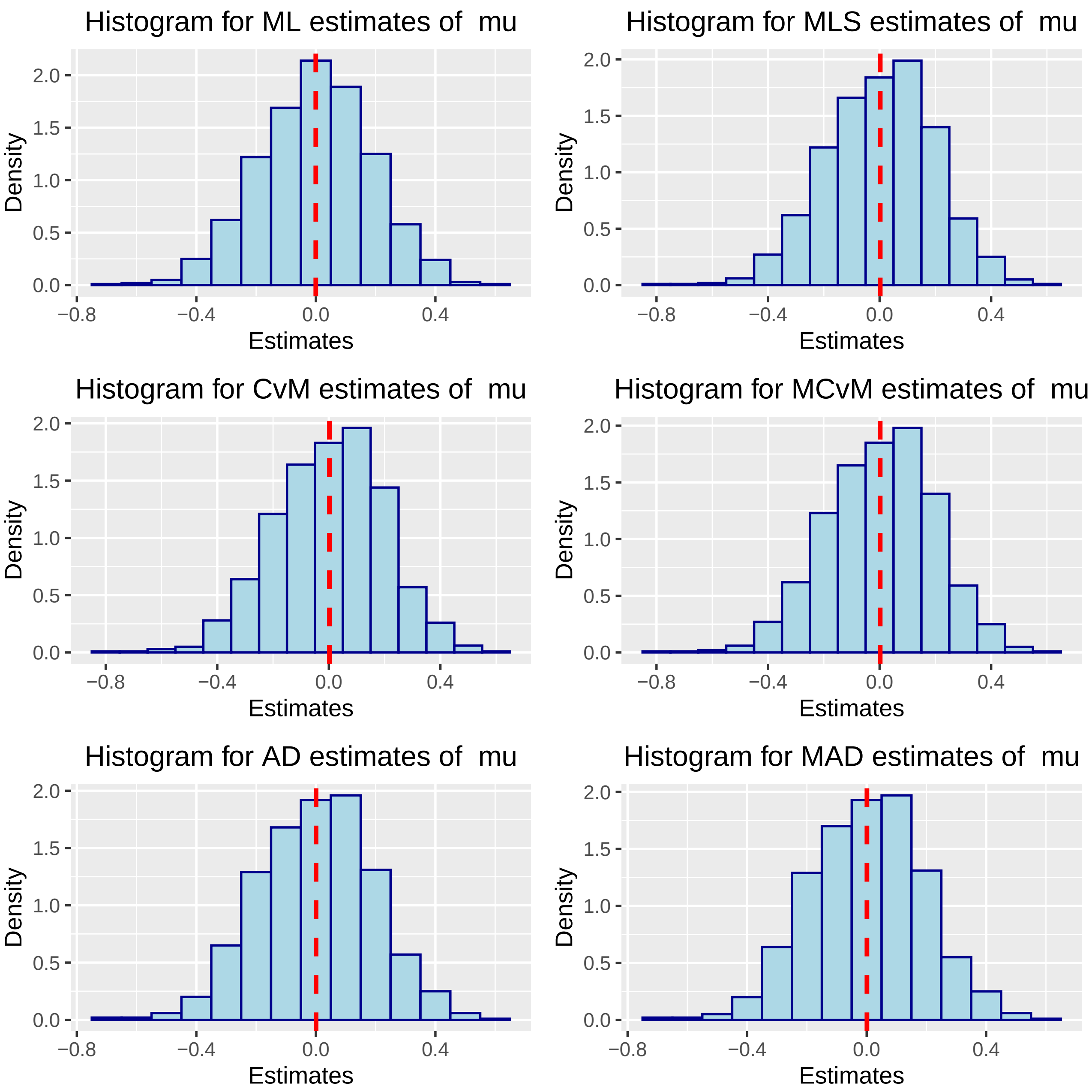
I. All of the estimators of μ have negligible bias. MLE has the least bias while MLSE has the maximum.

II. The MSE's of all the estimators of μ are low. MLE has the least MSE while CvM estimator has the maximum.

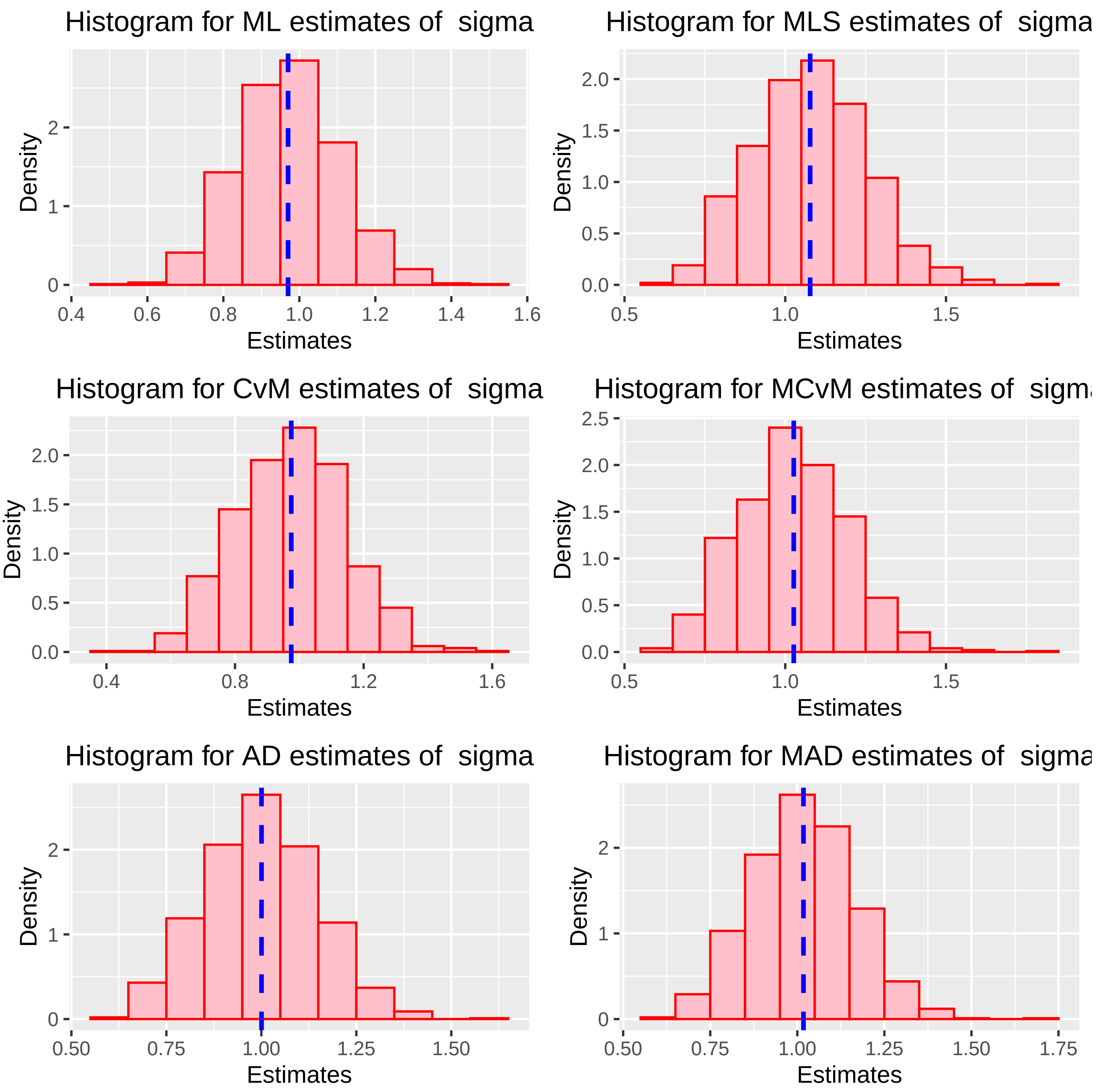
III. While considering bias of σ, AD estimator performs best and MLSE has the maximum bias. However, the bias of all other estimators are quite low.

IV. Again, for the case of MSE of σ, MLE outperforms the other estimators whereas MLSE performs worst. Here MCvM estimator performs slightly better than the CvM estimator.

V. As expected, MLE has the least Def value among all the estimators. Other estimators also have quite low Def values.


Histogram for ML estimates of mu
Histogram for MLS estimates of mu
Histogram for CvM estimates of mu
Histogram for MCvM estimates of mu
Histogram for AD estimates of mu
Histogram for MAD estimates of mu
Density
2.0
1.5
1.0
0.5
0.0
−0.8
−0.4
0.0
0.4
Estimates

$Figure : 1$


Histogram for ML estimates of sigma
Density
0
1
2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
Estimates
Histogram for MLS estimates of sigma
Density
0.0
0.5
1.0
1.5
2.0
0.5
1.0
1.5
Estimates
Histogram for CvM estimates of sigma
Density
0.0
0.5
1.0
1.5
2.0
0.4
0.8
1.2
1.6
Estimates
Histogram for MCvM estimates of sigm
Density
0.0
0.5
1.0
1.5
2.0
2.5
0.5
1.0
1.5
Estimates
Histogram for AD estimates of sigma
Density
0
1
2
0.50
0.75
1.00
1.25
1.50
Estimates
Histogram for MAD estimates of sigma
Density
0
1
2
0.50
0.75
1.00
1.25
1.50
1.75
Estimates

$Figure : 2$

- **Sample Size (n) = 50**

The following table shows the Bias and MSE of the different estimators of μ and σ and the corresponding Def values for our simulated sample −

```
[1] "Sample size =  50"
[1] "Observation matrix is "
     mu.bias mu.mse sigma.bias sigma.mse    def
MLE   0.0021 0.0194     0.0117    0.0102 0.0297
MLS   0.0004 0.0208    -0.0559    0.0204 0.0412
CvM   0.0005 0.0209     0.0076    0.0162 0.0371
MCvM  0.0005 0.0208    -0.0252    0.0170 0.0378
AD    0.0013 0.0198    -0.0077    0.0125 0.0323
MAD   0.0013 0.0198    -0.0193    0.0128 0.0326
```

$Table : 2$

Observations −

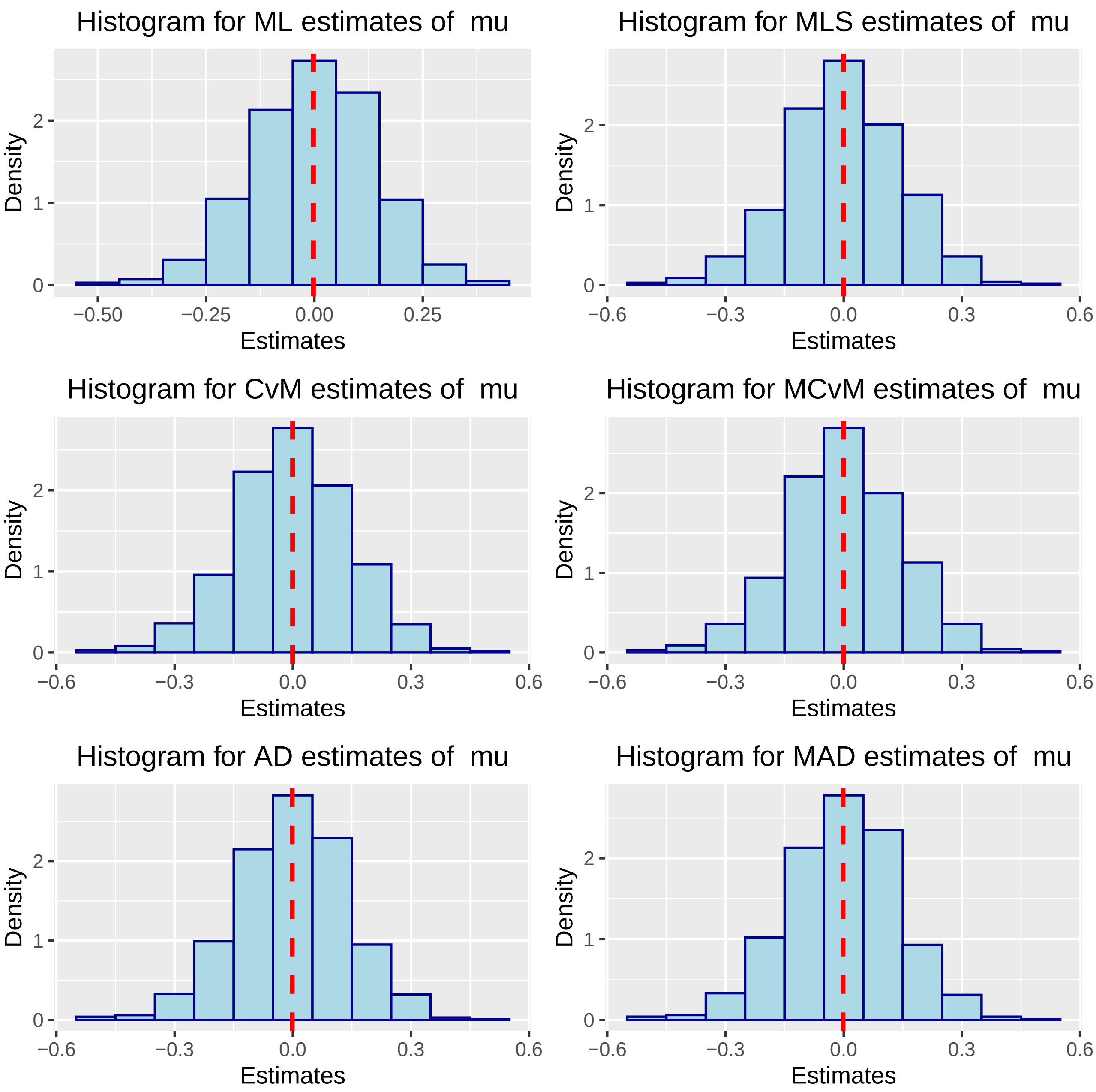
I. All of the estimators of μ have negligible bias. MLSE has the least bias while MLE has the maximum. Also CvM and MCvM are performing well.

II. The MSE's of all the estimators of μ are low. MLE has the least MSE while CvM estimator has the maximum. Here MCvM estimator performs slightly better than the CvM estimator.

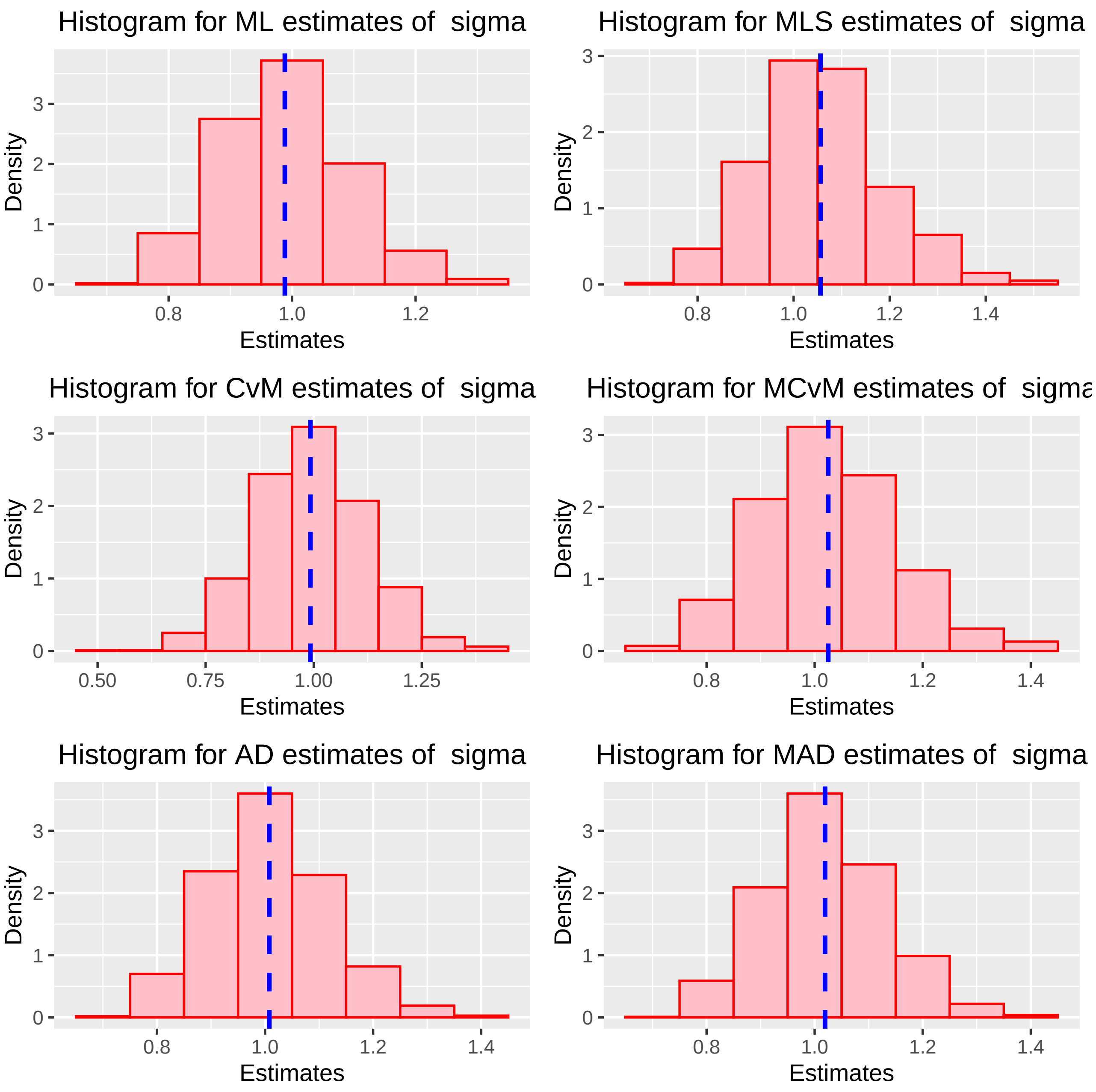
III. While considering bias of σ, CvM and AD estimators perform best and MLSE has the maximum bias.

IV. Again, for the case of MSE of σ, MLE outperforms the other estimators whereas MLSE performs worst.

V. As expected, MLE has the least Def value among all the estimators. Other estimators also have quite low Def values.


Histogram for ML estimates of mu
Density
0
1
2
−0.50
−0.25
0.00
0.25
Estimates
Histogram for MLS estimates of mu
Density
0
1
2
−0.6
−0.3
0.0
0.3
0.6
Estimates
Histogram for CvM estimates of mu
Density
0
1
2
−0.6
−0.3
0.0
0.3
0.6
Estimates
Histogram for MCvM estimates of mu
Density
0
1
2
−0.6
−0.3
0.0
0.3
0.6
Estimates
Histogram for AD estimates of mu
Density
0
1
2
−0.6
−0.3
0.0
0.3
0.6
Estimates
Histogram for MAD estimates of mu
Density
0
1
2
−0.6
−0.3
0.0
0.3
0.6
Estimates

$Figure : 3$



Histogram for ML estimates of sigma

Histogram for MLS estimates of sigma

Histogram for CvM estimates of sigma

Histogram for MCvM estimates of sigma

Histogram for AD estimates of sigma

Histogram for MAD estimates of sigma

$Figure : 4$

- **Sample Size (n) = 100**

The following table shows the Bias and MSE of the different estimators of μ and σ and the corresponding Def values for our simulated sample −

```
[1] "Sample size =  100"
[1] "Observation matrix is "
     mu.bias mu.mse sigma.bias sigma.mse    def
MLE  -0.0036 0.0096     0.0058    0.0049 0.0145
MLS  -0.0033 0.0103    -0.0282    0.0087 0.0189
CvM  -0.0033 0.0103     0.0035    0.0076 0.0179
MCvM -0.0033 0.0103    -0.0130    0.0078 0.0181
AD   -0.0034 0.0098    -0.0036    0.0059 0.0156
MAD  -0.0034 0.0098    -0.0099    0.0059 0.0157
```

$Table : 3$

Observations −

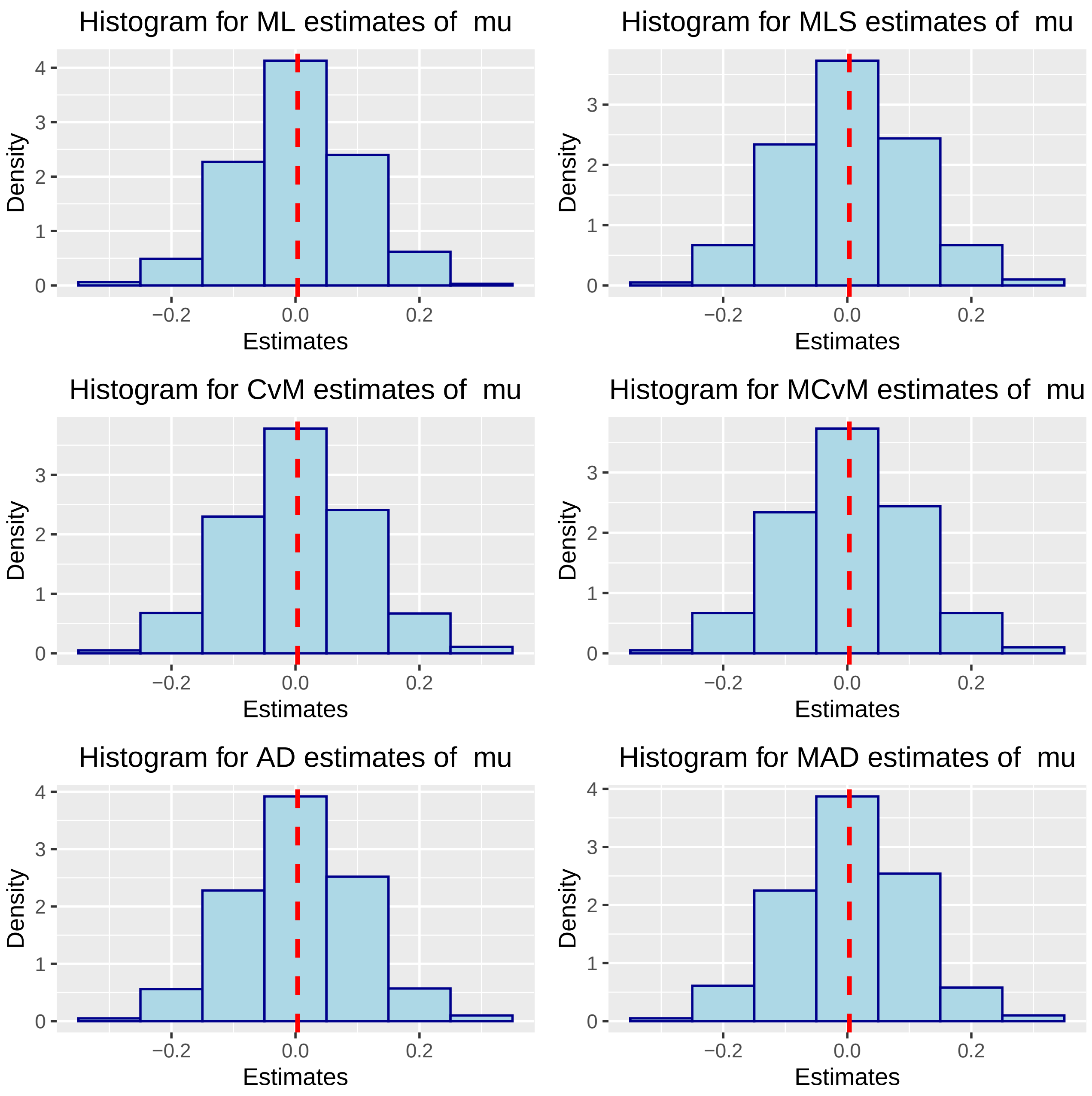
I. All of the estimators of μ have negligible bias.

II. The MSE's of all the estimators of μ are low.

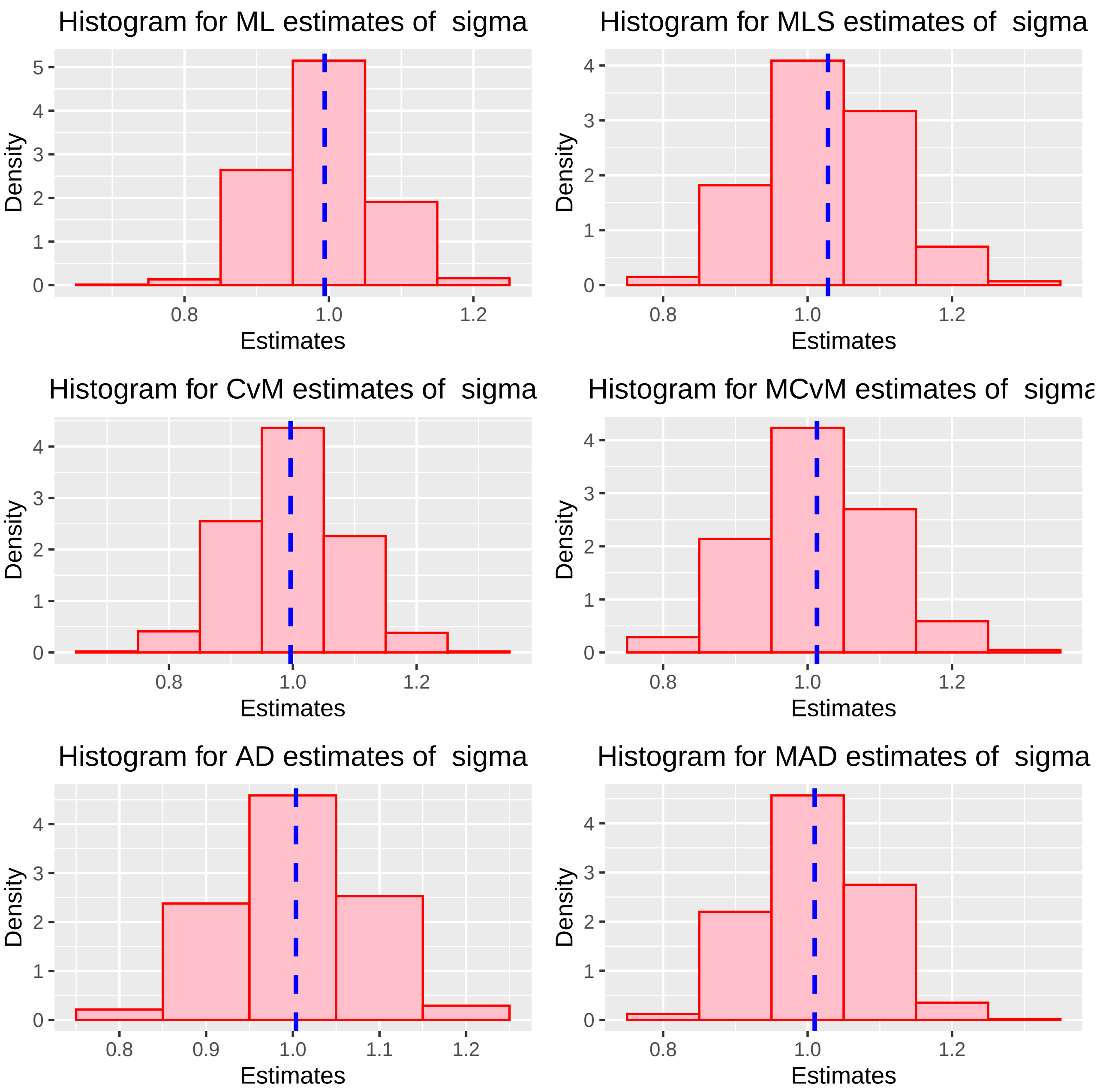
III. While considering bias of σ, CvM and AD estimators perform best and MLSE has the maximum bias.

IV. Again, for the case of MSE of σ, MLE outperforms the other estimators whereas MLSE performs worst.

V. As expected, MLE has the least Def value among all the estimators. Other estimators also have quite low Def values.


Histogram for ML estimates of mu
Histogram for MLS estimates of mu
Histogram for CvM estimates of mu
Histogram for MCvM estimates of mu
Histogram for AD estimates of mu
Histogram for MAD estimates of mu
Density
Estimates
0
1
2
3
4
−0.2
0.0
0.2

$Figure : 5$



Histogram for ML estimates of sigma

Histogram for MLS estimates of sigma

Histogram for CvM estimates of sigma

Histogram for MCvM estimates of sigma

Histogram for AD estimates of sigma

Histogram for MAD estimates of sigma

Figure : 6

NOTE : It can be observed that as the sample size increases, the MCvM and MAD estimators of the location parameter μ and scale parameter σ are performing very close to their counterparts CvM and AD estimators.

4.1.2 Sampling from Extreme Value (0,1)

We will now show the results of our simulation study for the different sample sizes :

- **Sample Size (n) = 30**

The following table shows the Bias and MSE of the different estimators of μ and σ and the corresponding Def values for our simulated sample −

```
[1] "Sample size =  30"
[1] "Observation matrix is "
     mu.bias mu.mse sigma.bias sigma.mse    def
MLE  -0.0220 0.0372     0.0266    0.0198 0.0569
MLS  -0.0002 0.0404    -0.0901    0.0409 0.0814
CvM  -0.0228 0.0416     0.0172    0.0303 0.0719
MCvM -0.0154 0.0408    -0.0382    0.0314 0.0722
AD   -0.0158 0.0384    -0.0089    0.0232 0.0616
MAD  -0.0227 0.0388    -0.0298    0.0244 0.0632
```

$Table : 4$

Observations −

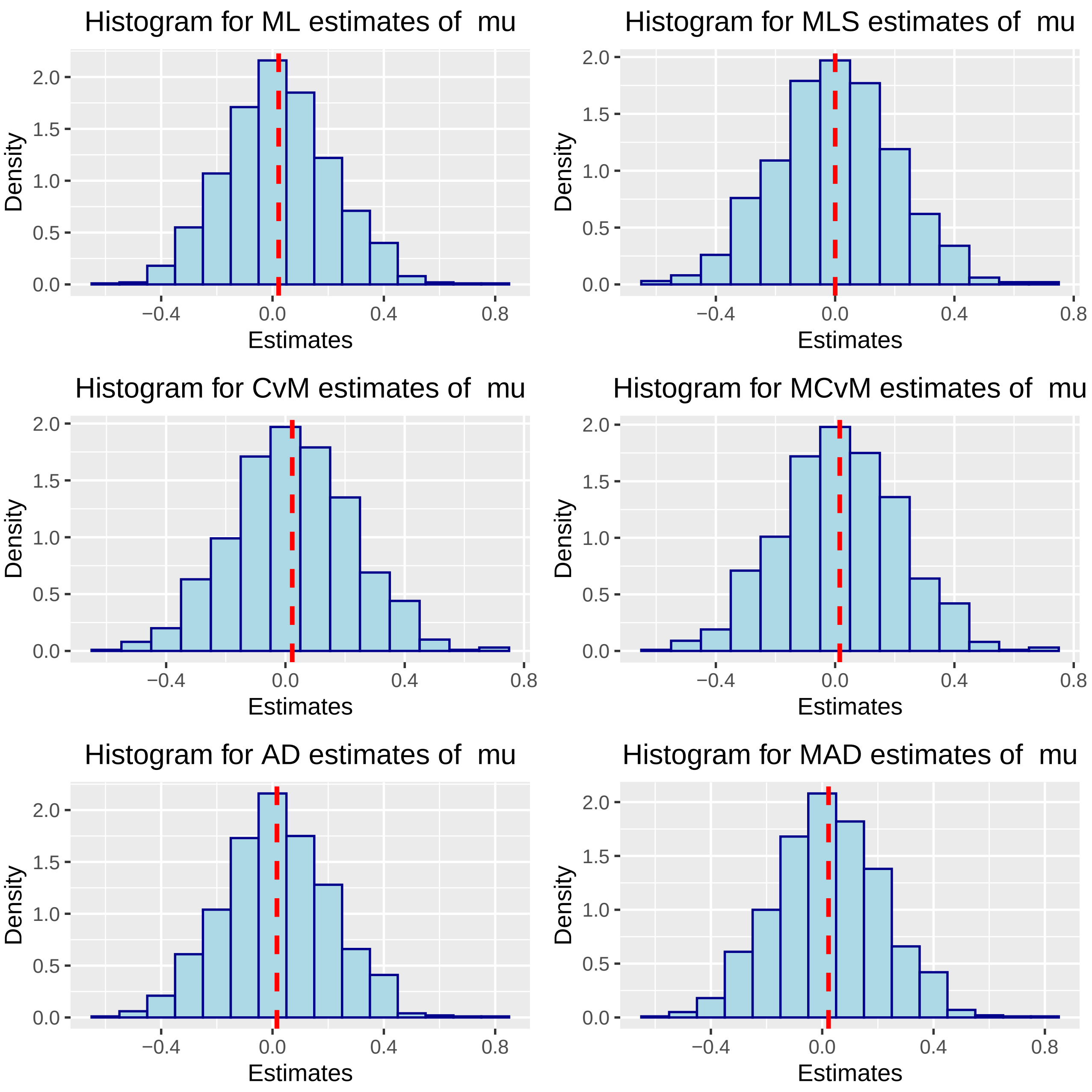
I. MLSE has the least bias. It's value is significantly lower than the other estimators.

II. The MSE's of all the estimators of μ are low.

III. While considering bias of σ, AD estimator perform best and MLSE has the maximum bias.

IV. For the case of MSE of σ, MLE outperforms the other estimators whereas MLSE performs worst.

V. MLE has the least Def value among all the estimators. MLSE has the highest Def value.



Histogram for ML estimates of mu

Histogram for MLS estimates of mu

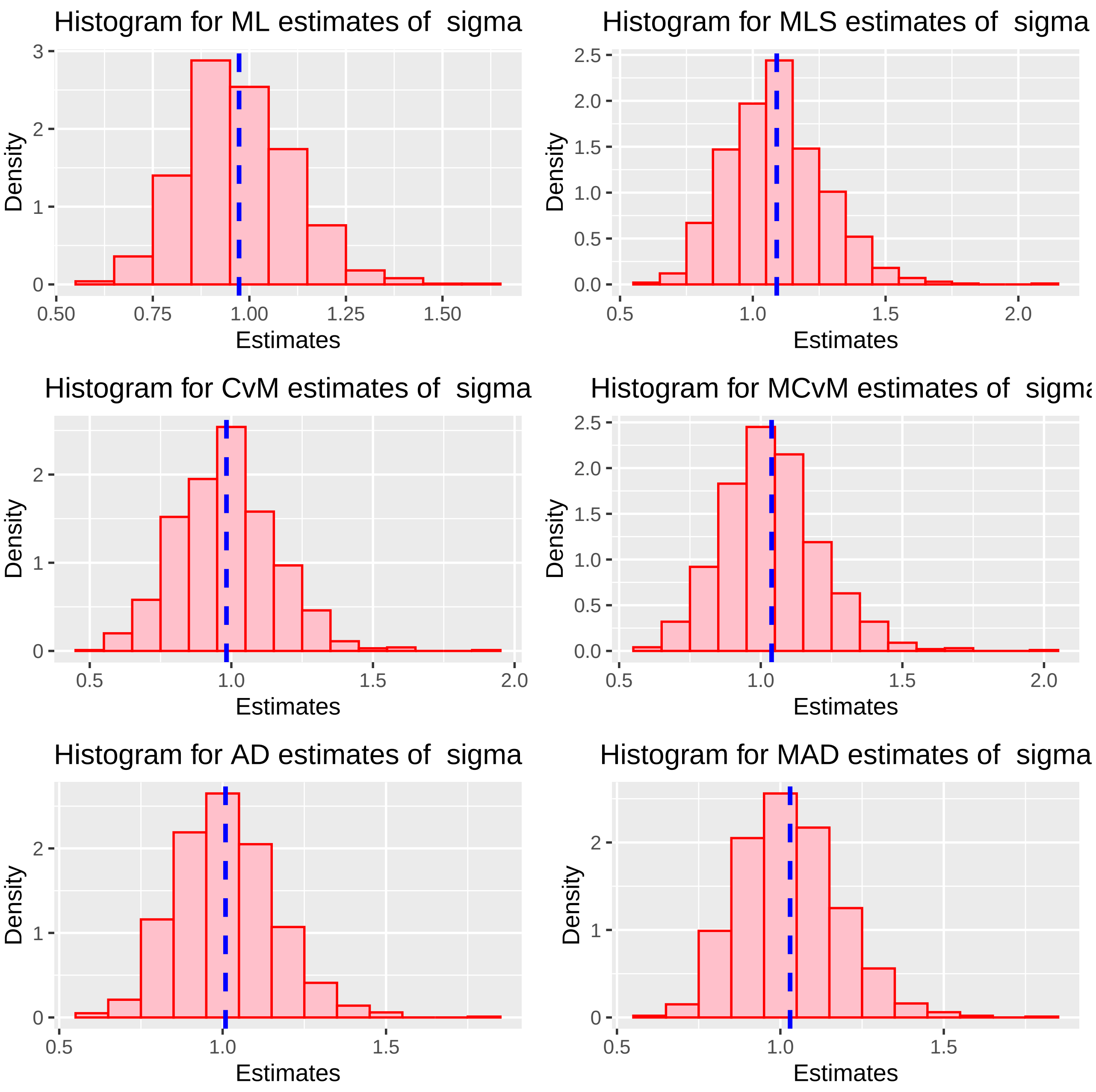
Histogram for CvM estimates of mu

Histogram for MCvM estimates of mu

Histogram for AD estimates of mu

Histogram for MAD estimates of mu

$Figure : 7$



Histogram for ML estimates of sigma

Histogram for MLS estimates of sigma

Histogram for CvM estimates of sigma

Histogram for MCvM estimates of sigma

Histogram for AD estimates of sigma

Histogram for MAD estimates of sigma

Figure : 8

- **Sample Size (n) = 50**

The following table shows the Bias and MSE of the different estimators of μ and σ and the corresponding Def values for our simulated sample −

```
[1] "Sample size =  50"
[1] "Observation matrix is "
     mu.bias mu.mse sigma.bias sigma.mse    def
MLE  -0.0057 0.0208     0.0145    0.0130 0.0338
MLS   0.0064 0.0234    -0.0564    0.0234 0.0468
CvM  -0.0073 0.0238     0.0101    0.0191 0.0429
MCvM -0.0026 0.0235    -0.0256    0.0198 0.0432
AD   -0.0024 0.0217    -0.0059    0.0150 0.0367
MAD  -0.0068 0.0218    -0.0206    0.0156 0.0374
```

$Table : 5$

Observations −

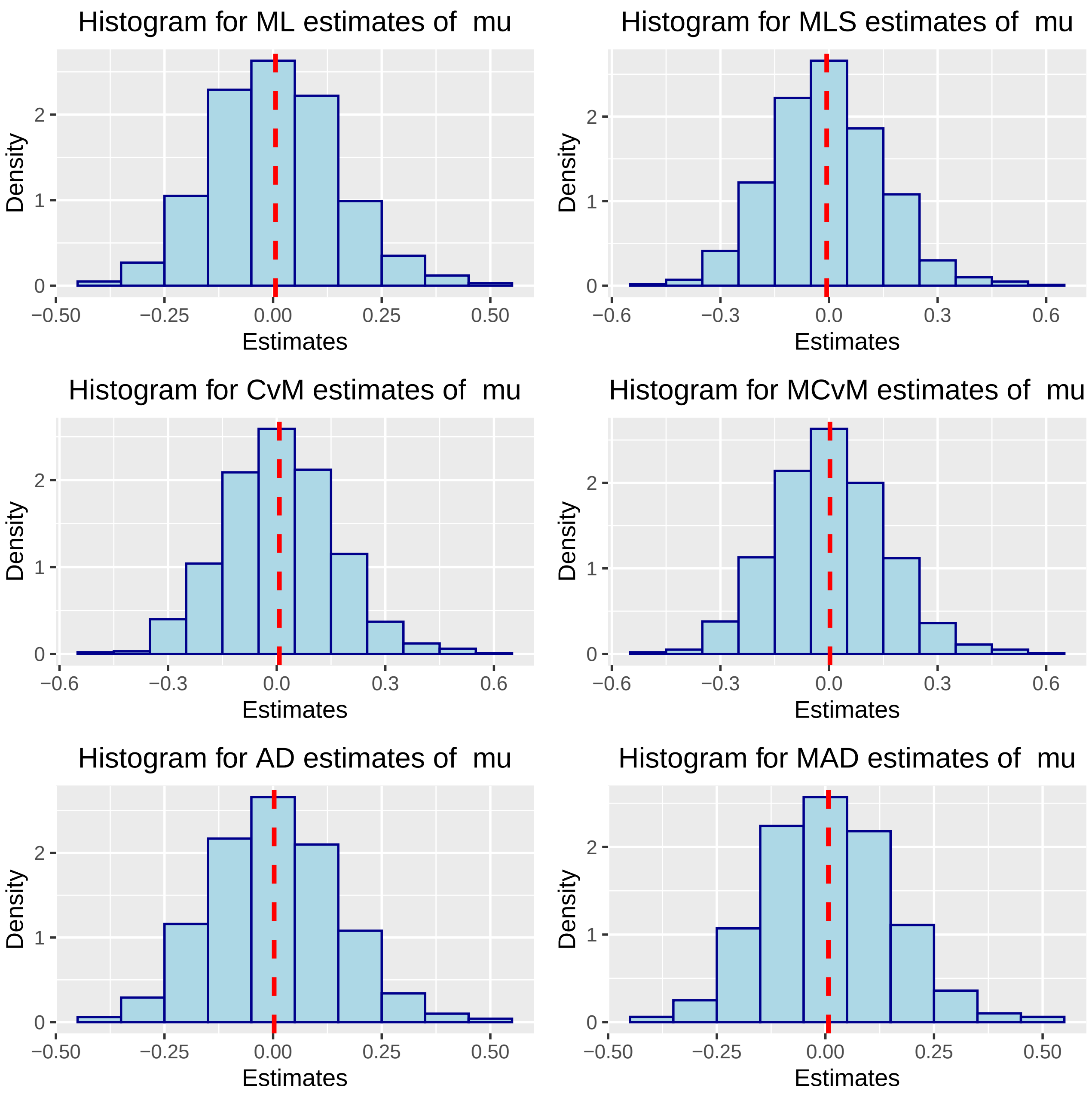
I. AD and MCvM estiamtors have the least biases. CvM estiamtor has the highest bias. Here MCvM performs better than CvM.

II. The MSE's of all the estimators of μ are low and quite similar to each other.

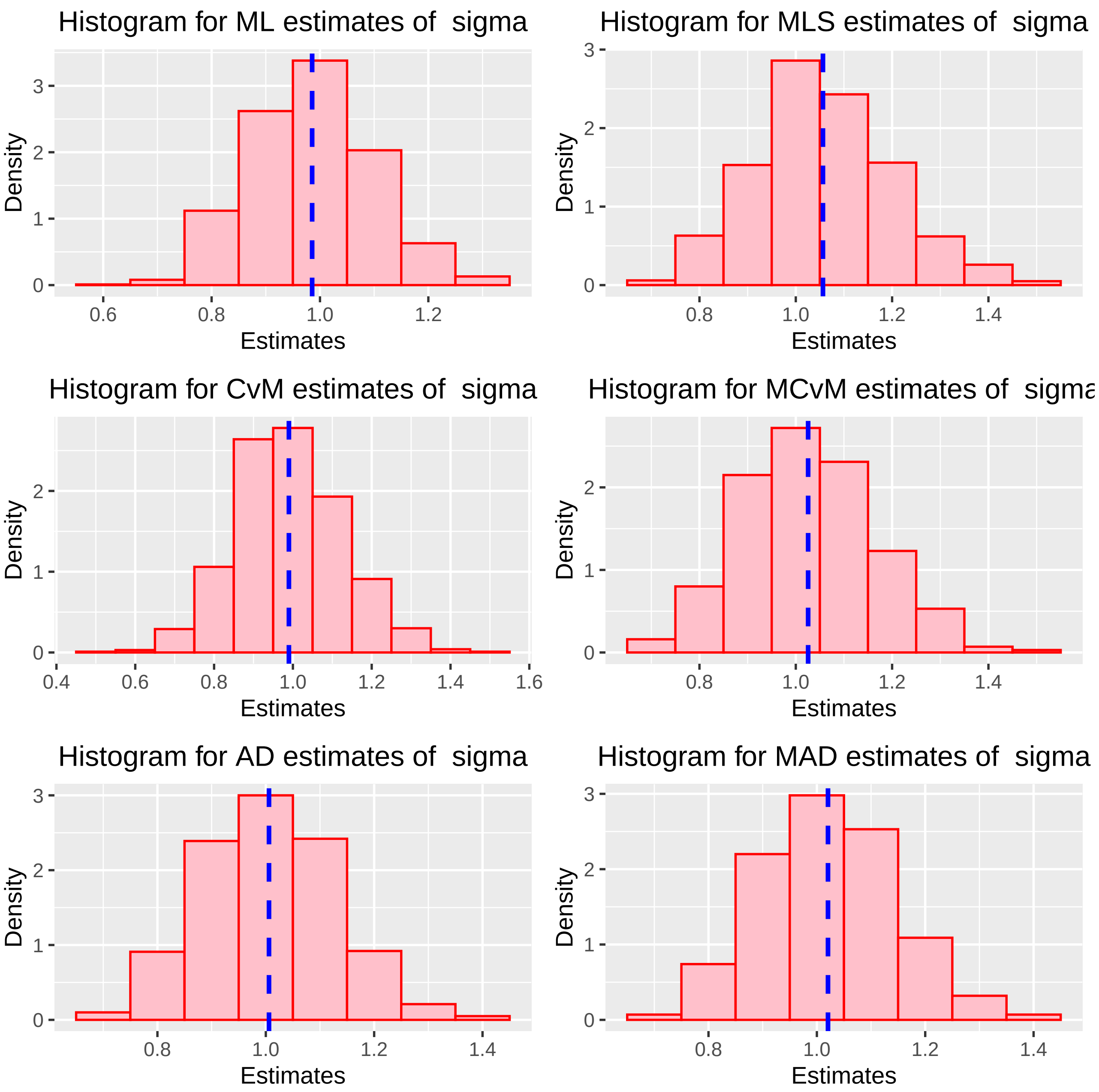
III. While considering bias of σ, AD estimator performs best and MLSE has the maximum bias.

IV. For the case of MSE of σ, MLE outperforms the other estimators whereas MLSE performs worst.

V. MLE has the least Def value among all the estimators. MLSE has the highest Def value.


Histogram for ML estimates of mu
Histogram for MLS estimates of mu
Histogram for CvM estimates of mu
Histogram for MCvM estimates of mu
Histogram for AD estimates of mu
Histogram for MAD estimates of mu
Density
Estimates
0
1
2
−0.50
−0.25
0.00
0.25
0.50
−0.6
−0.3
0.0
0.3
0.6

$Figure : 9$


Histogram for ML estimates of sigma
Density
Estimates
0
1
2
3
0.6
0.8
1.0
1.2
Histogram for MLS estimates of sigma
Density
Estimates
0
1
2
3
0.8
1.0
1.2
1.4
Histogram for CvM estimates of sigma
Density
Estimates
0
1
2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
Histogram for MCvM estimates of sigma
Density
Estimates
0
1
2
0.8
1.0
1.2
1.4
Histogram for AD estimates of sigma
Density
Estimates
0
1
2
3
0.8
1.0
1.2
1.4
Histogram for MAD estimates of sigma
Density
Estimates
0
1
2
3
0.8
1.0
1.2
1.4

$Figure : 10$

- **Sample Size (n) = 100**

The following table shows the Bias and MSE of the different estimators of μ and σ and the corresponding Def values for our simulated sample −

```
[1] "Sample size =  100"
[1] "Observation matrix is "
     mu.bias mu.mse sigma.bias sigma.mse    def
MLE   0.0010 0.0106     0.0118    0.0062 0.0168
MLS   0.0083 0.0114    -0.0226    0.0098 0.0211
CvM   0.0015 0.0114     0.0101    0.0091 0.0205
MCvM  0.0038 0.0113    -0.0075    0.0091 0.0204
AD    0.0034 0.0107     0.0018    0.0071 0.0178
MAD   0.0014 0.0107    -0.0061    0.0072 0.0179
```

$Table : 6$

Observations −

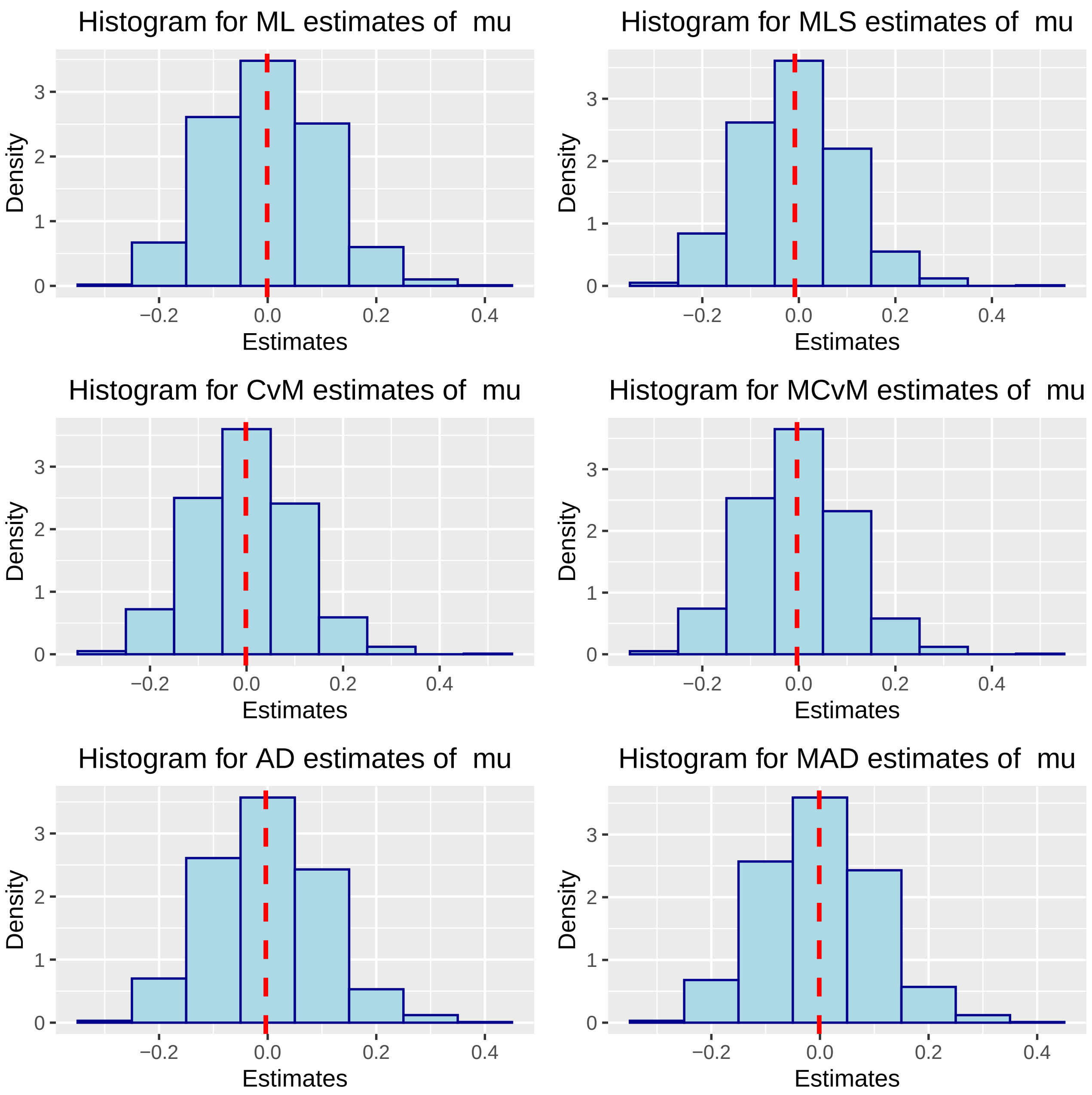
I. MLE have the least biases. MLSE estiamtor has the highest bias. Here MAD estimators performs better than AD estimators.

II. The MSE's of all the estimators of μ are low and quite similar to each other.

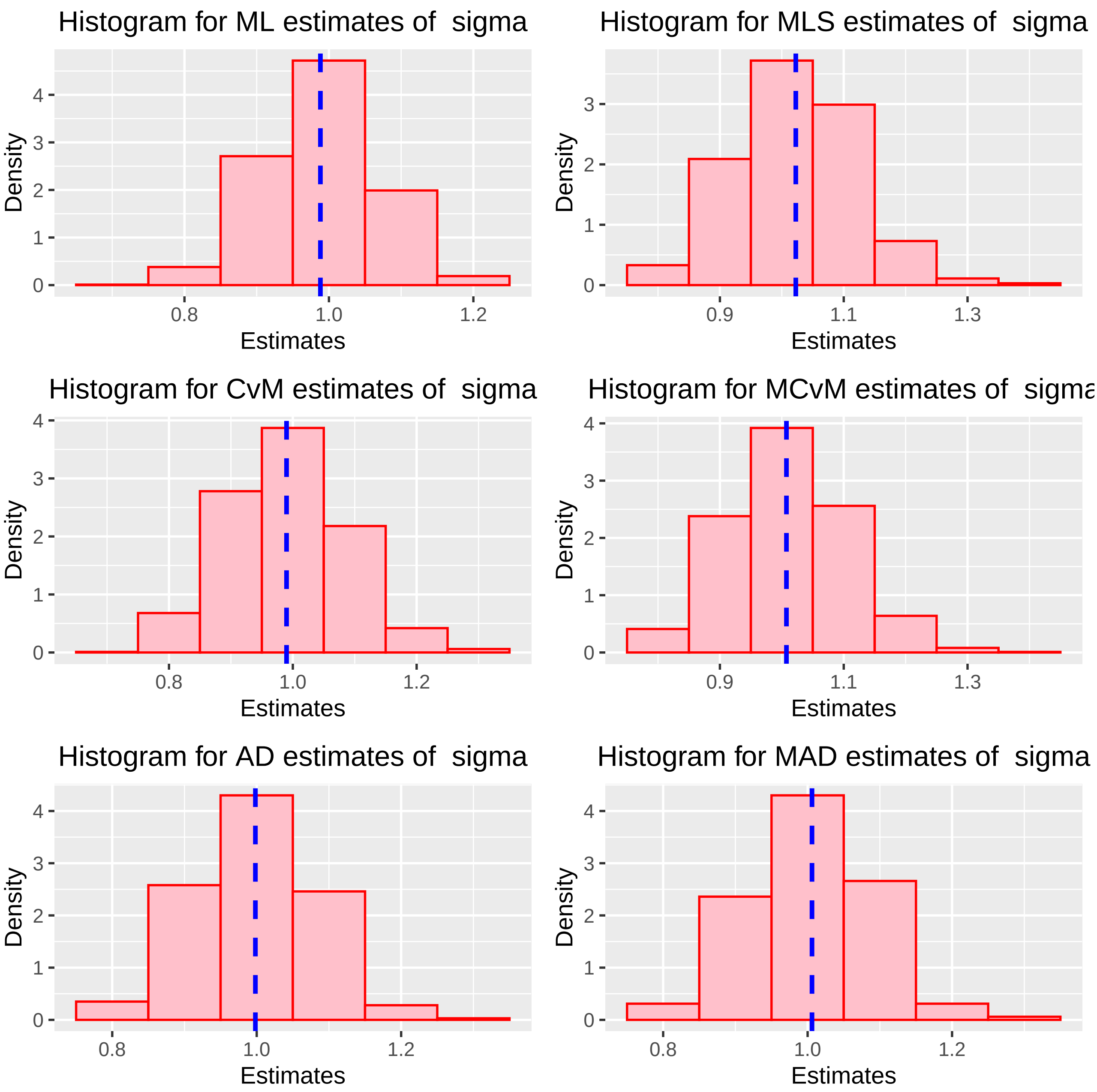
III. While considering bias of σ, AD estimator performs best and MLSE has the maximum bias.

IV. For the case of MSE of σ, MLE outperforms the other estimators whereas MLSE performs worst.

V. MLE has the least Def value among all the estimators. MLSE has the highest Def value.


Histogram for ML estimates of mu
Histogram for MLS estimates of mu
Histogram for CvM estimates of mu
Histogram for MCvM estimates of mu
Histogram for AD estimates of mu
Histogram for MAD estimates of mu
Density
0
1
2
3
Estimates
−0.2
0.0
0.2
0.4

$Figure : 11$



Histogram for ML estimates of sigma

Histogram for MLS estimates of sigma

Histogram for CvM estimates of sigma

Histogram for MCvM estimates of sigma

Histogram for AD estimates of sigma

Histogram for MAD estimates of sigma

$Figure$: 12

NOTE : It can be observed that as the sample size increases, the MCvM and MAD estimators of the location parameter μ and scale parameter σ are performing very close to their counterparts CvM and AD estimators.

4.1.3 Sampling from Weibull (0,1,1.2)

We will now show the results of our simulation study for the different sample sizes :

- **Sample Size (n) = 30**

The following table shows the Bias and MSE of the different estimators of μ and σ and the corresponding Def values for our simulated sample −

```
[1] "Sample size =  30"
[1] "Observation matrix is "
     mu.bias mu.mse sigma.bias sigma.mse    def
MLE  -0.0458 0.0040     0.0486    0.0246 0.0286
MLS   0.0291 0.0078    -0.0721    0.0416 0.0494
CvM  -0.0236 0.0104     0.0249    0.0361 0.0465
MCvM  0.0014 0.0066    -0.0291    0.0345 0.0411
AD    0.0053 0.0034    -0.0159    0.0286 0.0319
MAD  -0.0098 0.0042    -0.0182    0.0282 0.0324
```

$Table : 7$

Observations −

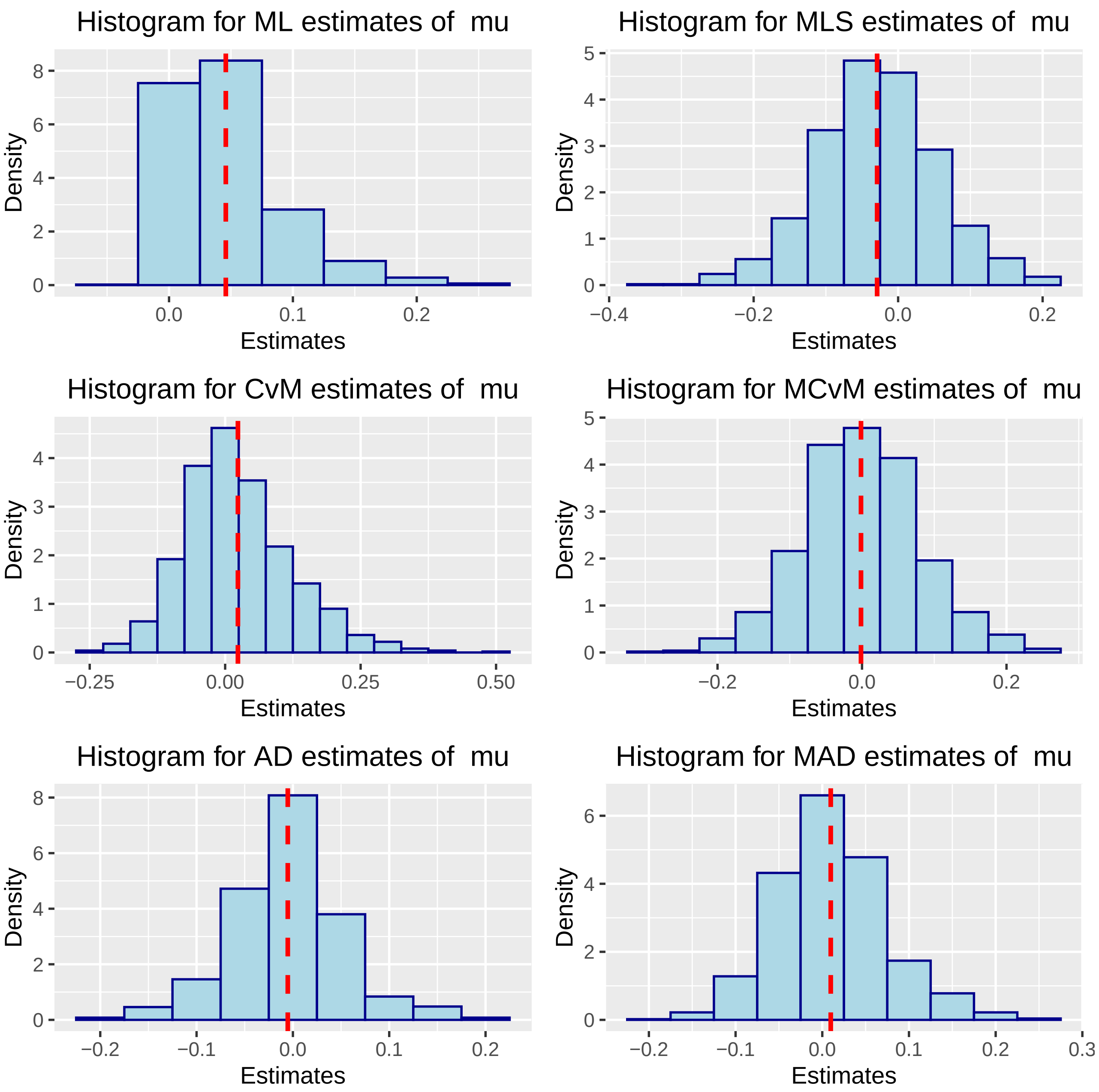
I. MCvM has the least bias. It's value is significantly lower than the other estimators. MLE performs worst in this case.

II. The MSE's of all the estimators of μ are low. AD estimator has the lowest MSE and CvM estimator has the highest value. Here MCvM estimator performs better that CvM estimator.

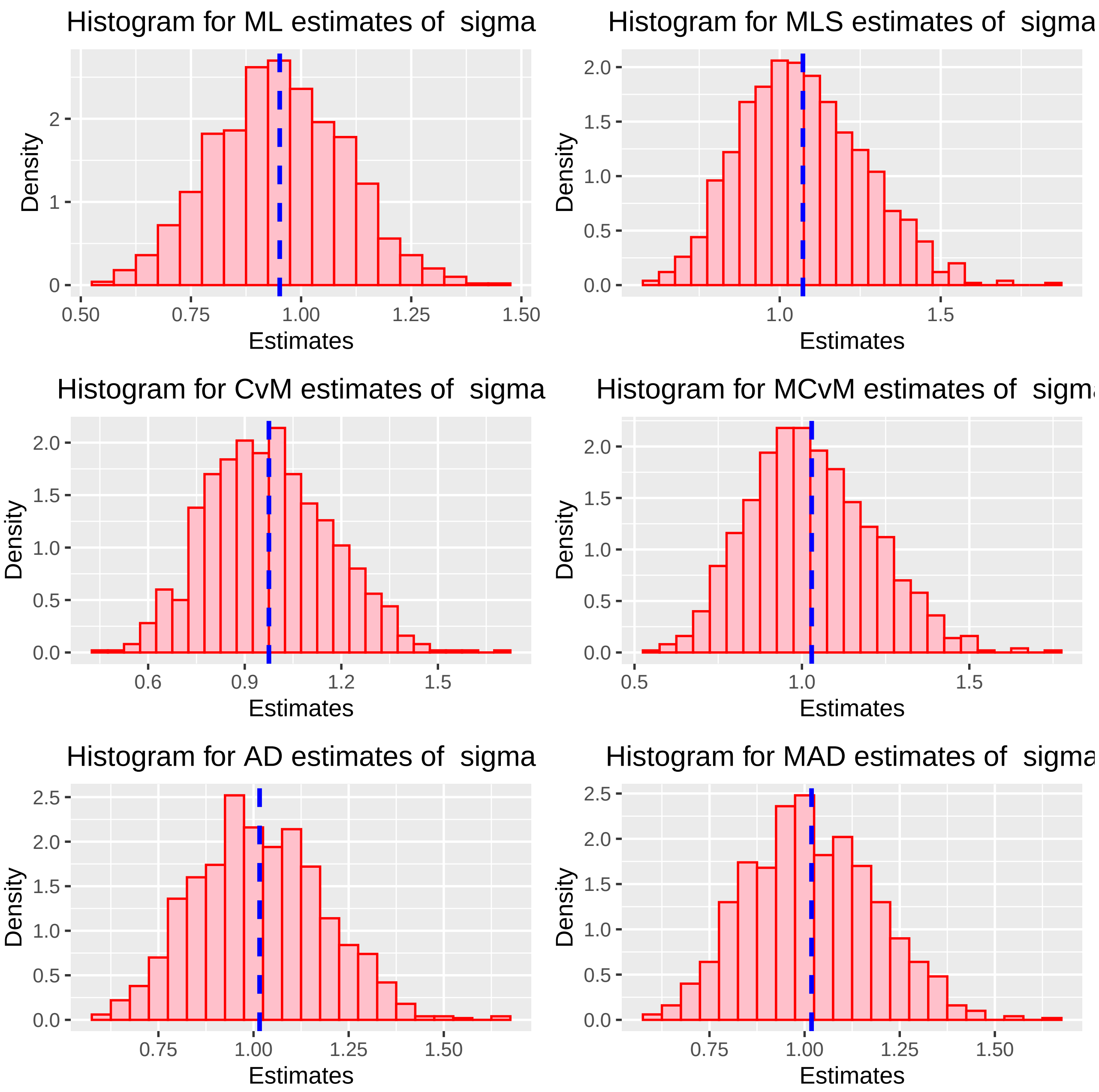
III. While considering bias of σ, AD estimator performs best and MLSE has the maximum bias.

IV. For the case of MSE of σ, MLE outperforms the other estimators whereas MLSE performs worst. Also MCvM and MAD estimators outperform CvM and AD estimators respectively.

V. MLE has the least Def value among all the estimators. MLSE has the highest Def value.


Histogram for ML estimates of mu
Density
8
6
4
2
0
0.0
0.1
0.2
Estimates
Histogram for MLS estimates of mu
Density
5
4
3
2
1
0
−0.4
−0.2
0.0
0.2
Estimates
Histogram for CvM estimates of mu
Density
4
3
2
1
0
−0.25
0.00
0.25
0.50
Estimates
Histogram for MCvM estimates of mu
Density
5
4
3
2
1
0
−0.2
0.0
0.2
Estimates
Histogram for AD estimates of mu
Density
8
6
4
2
0
−0.2
−0.1
0.0
0.1
0.2
Estimates
Histogram for MAD estimates of mu
Density
6
4
2
0
−0.2
−0.1
0.0
0.1
0.2
0.3
Estimates

$Figure : 13$



Histogram for ML estimates of sigma

Histogram for MLS estimates of sigma

Histogram for CvM estimates of sigma

Histogram for MCvM estimates of sigm

Histogram for AD estimates of sigma

Histogram for MAD estimates of sigma

Figure : 14

- **Sample Size (n) = 50**

The following table shows the Bias and MSE of the different estimators of μ and σ and the corresponding Def values for our simulated sample −

```
[1] "Sample size =  50"
[1] "Observation matrix is "
     mu.bias mu.mse sigma.bias sigma.mse    def
MLE  -0.0306 0.0017     0.0354    0.0154 0.0171
MLS   0.0189 0.0044    -0.0403    0.0233 0.0277
CvM  -0.0123 0.0053     0.0181    0.0217 0.0270
MCvM  0.0024 0.0039    -0.0147    0.0209 0.0248
AD    0.0067 0.0016    -0.0101    0.0174 0.0190
MAD  -0.0049 0.0021    -0.0090    0.0172 0.0193
```

$Table : 8$

Observations −

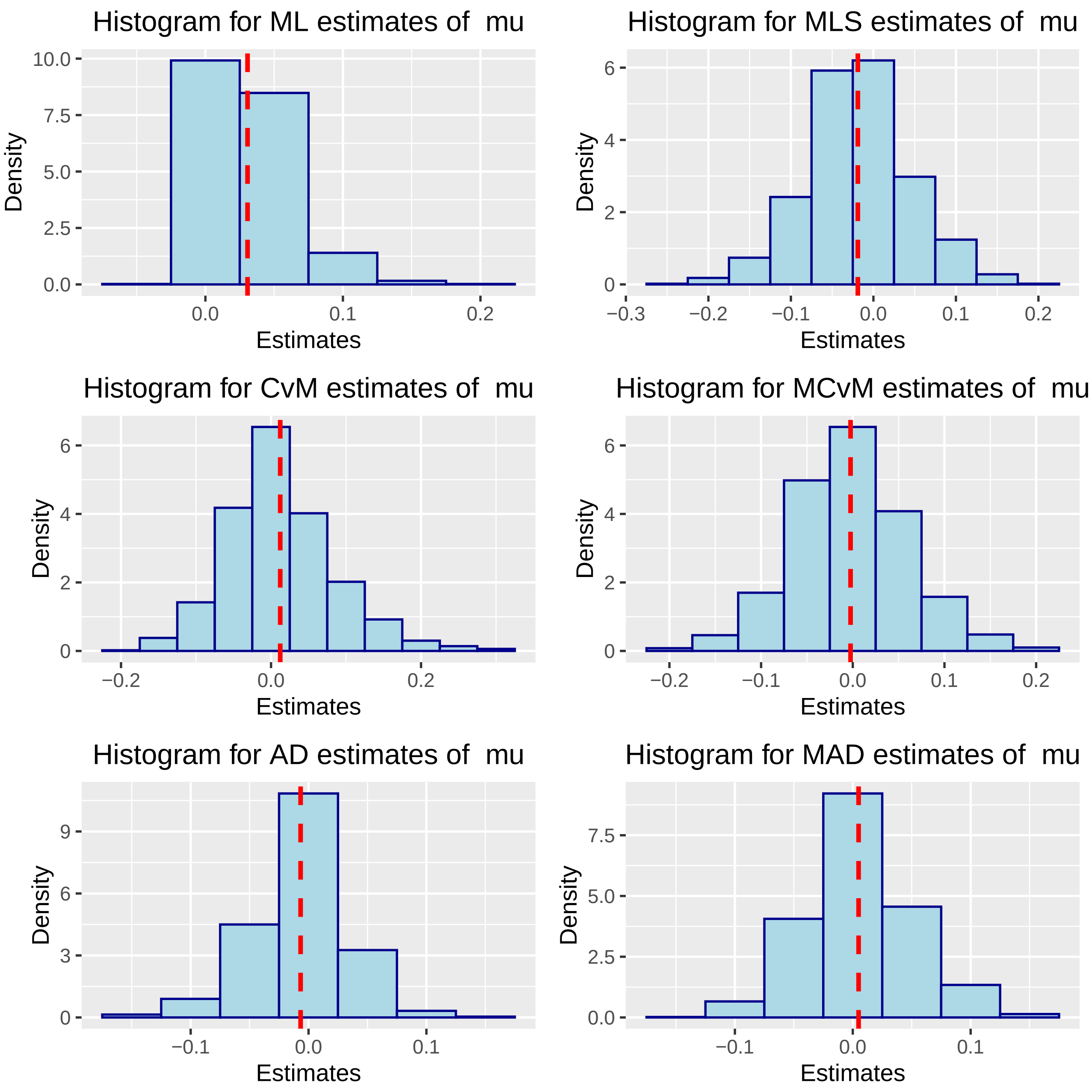
I. MCvM has the least bias. It's value is significantly lower than the other estimators. MLE performs worst in this case.

II. The MSE's of all the estimators of μ are low. AD estimator has the lowest MSE and CvM estimator has the highest value. Here MCvM estimator performs better that CvM estimator.

III. While considering bias of σ, MAD estimator perform best and MLSE has the maximum bias.

IV. For the case of MSE of σ, MLE outperforms the other estimators whereas MLSE performs worst. Also MCvM and MAD estimators outperform CvM and AD estimators respectively.

V. MLE has the least Def value among all the estimators. MLSE has the highest Def value.



Histogram for ML estimates of mu

Histogram for MLS estimates of mu

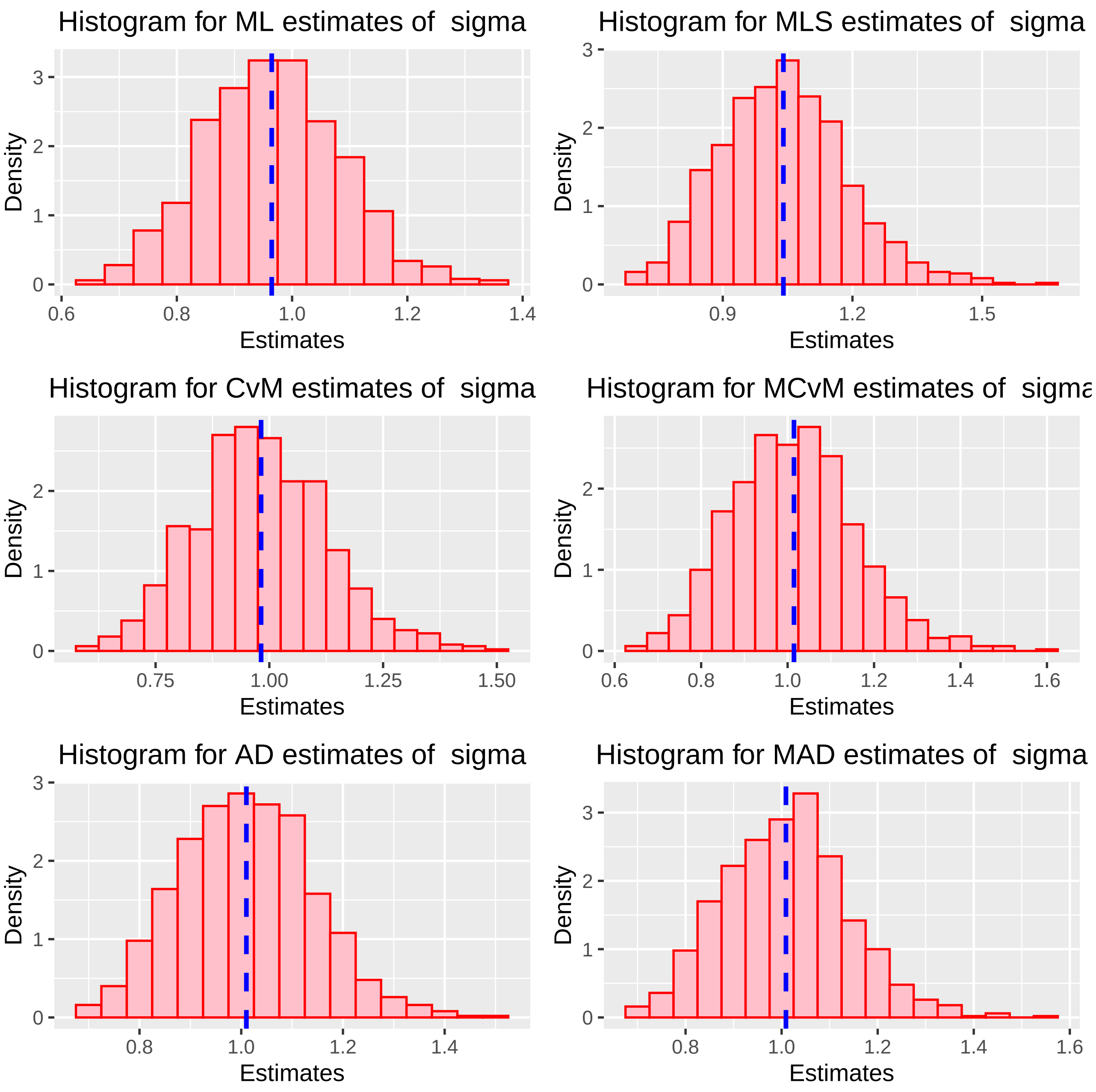
Histogram for CvM estimates of mu

Histogram for MCvM estimates of mu

Histogram for AD estimates of mu

Histogram for MAD estimates of mu

$Figure : 15$


Histogram for ML estimates of sigma
Density
Estimates
0
1
2
3
0.6
0.8
1.0
1.2
1.4
Histogram for MLS estimates of sigma
Density
Estimates
0
1
2
3
0.9
1.2
1.5
Histogram for CvM estimates of sigma
Density
Estimates
0
1
2
0.75
1.00
1.25
1.50
Histogram for MCvM estimates of sigma
Density
Estimates
0
1
2
0.6
0.8
1.0
1.2
1.4
1.6
Histogram for AD estimates of sigma
Density
Estimates
0
1
2
3
0.8
1.0
1.2
1.4
Histogram for MAD estimates of sigma
Density
Estimates
0
1
2
3
0.8
1.0
1.2
1.4
1.6

$Figure : 16$

- **Sample Size (n) = 100**

The following table shows the Bias and MSE of the different estimators of μ and σ and the corresponding Def values for our simulated sample −

```
[1] "Sample size =  100"
[1] "Observation matrix is "
     mu.bias mu.mse sigma.bias sigma.mse    def
MLE  -0.0194 0.0007     0.0179    0.0075 0.0082
MLS   0.0092 0.0024    -0.0223    0.0118 0.0141
CvM  -0.0061 0.0026     0.0062    0.0112 0.0138
MCvM  0.0010 0.0023    -0.0096    0.0111 0.0134
AD    0.0056 0.0008    -0.0112    0.0088 0.0096
MAD  -0.0030 0.0010    -0.0070    0.0088 0.0098
```

$Table : 9$

Observations −

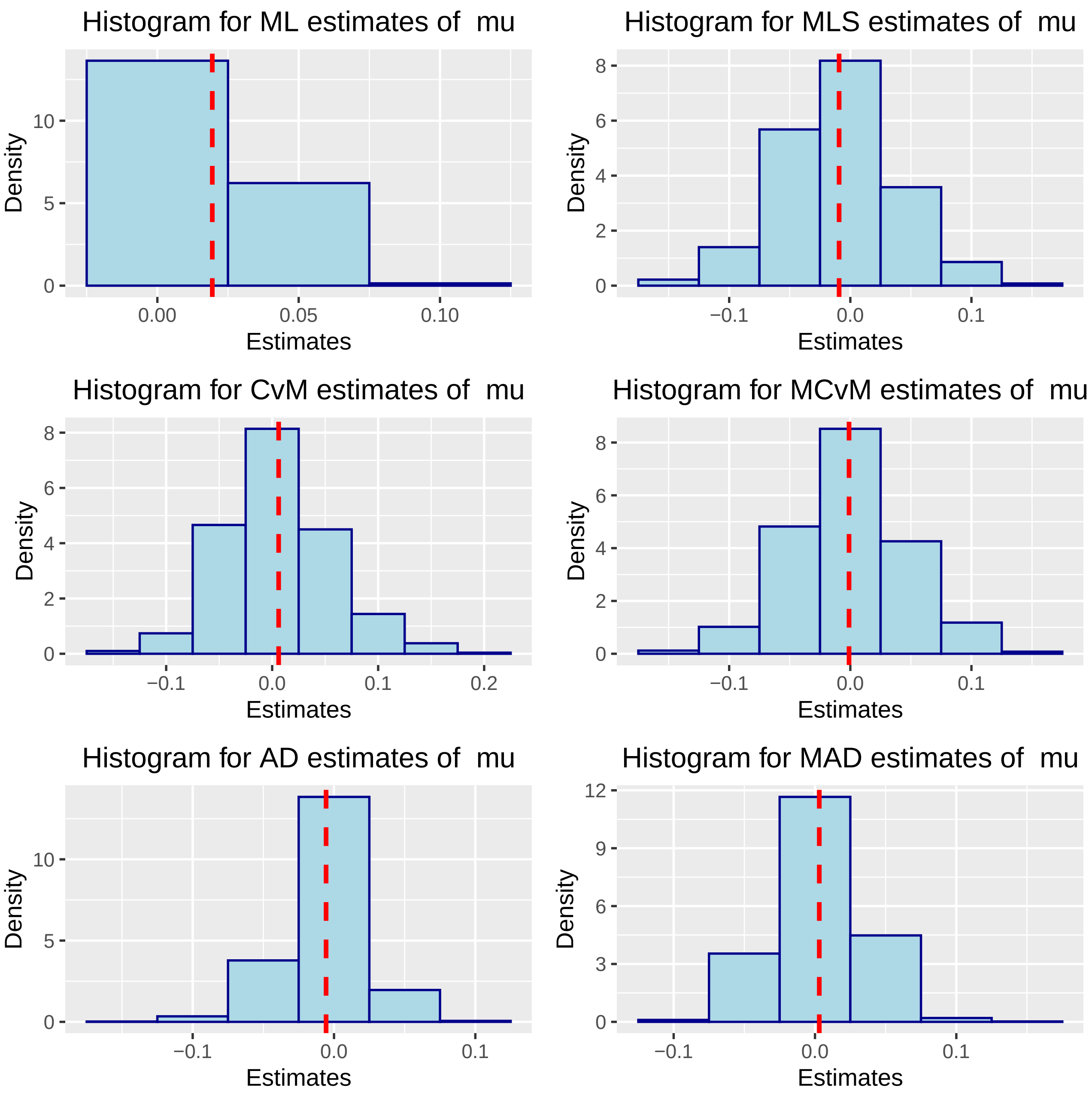
I. MCvM has the least bias. It's value is significantly lower than the other estimators. MLE performs worst in this case.

II. The MSE's of all the estimators of μ are low. MLE,AD,MAD estimators have the lowest MSE's and CvM estimator has the highest value. Here MCvM estimator performs better that CvM estimator.

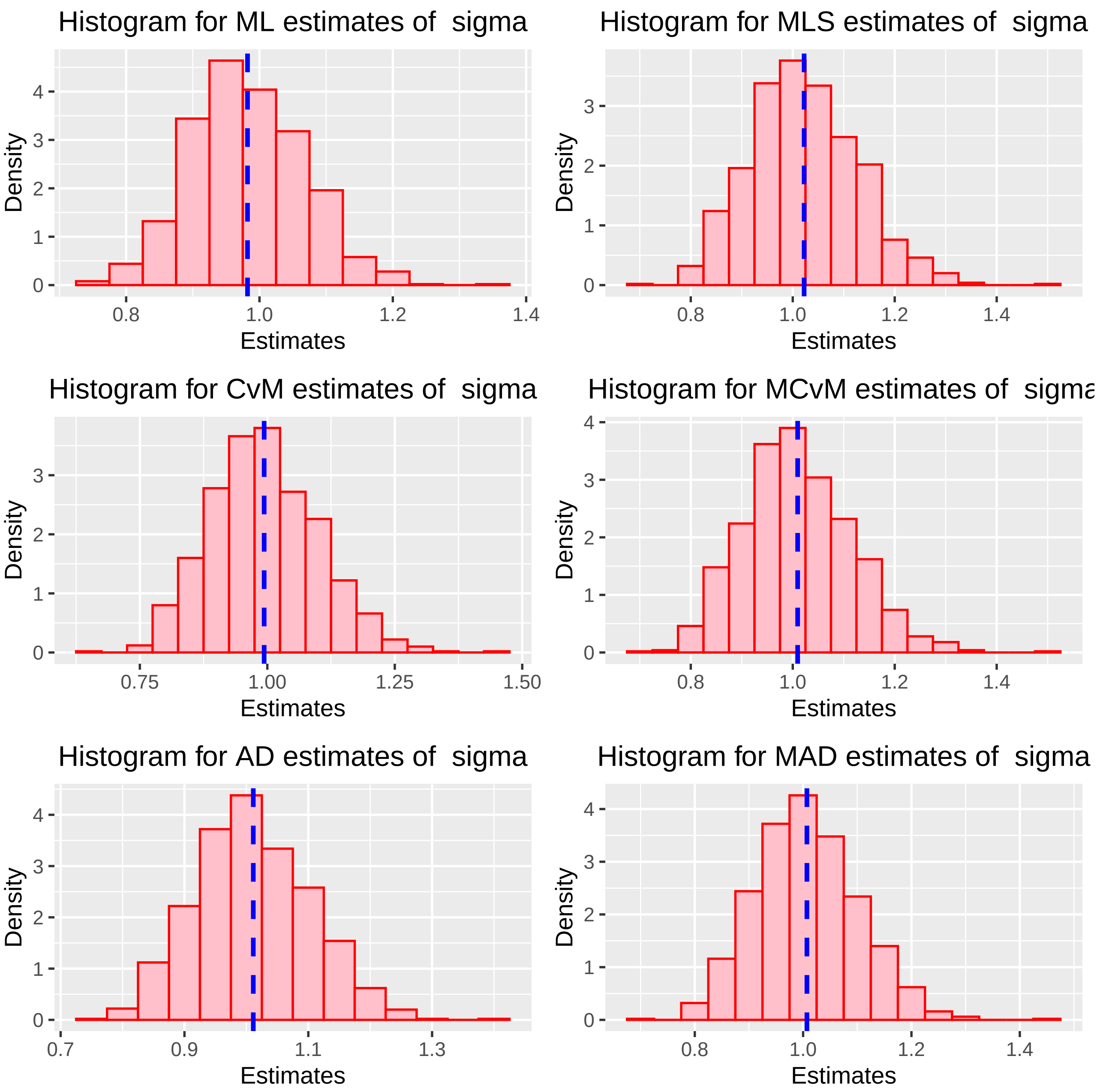
III. While considering bias of σ, CvM estimator perform best and MLSE has the maximum bias.

IV. For the case of MSE of σ, MLE outperforms the other estimators whereas MLSE performs worst.

V. MLE has the least Def value among all the estimators. MLSE has the highest Def value.


Histogram for ML estimates of mu
Density
10
5
0
0.00
0.05
0.10
Estimates
Histogram for MLS estimates of mu
Density
8
6
4
2
0
−0.1
0.0
0.1
Estimates
Histogram for CvM estimates of mu
Density
8
6
4
2
0
−0.1
0.0
0.1
0.2
Estimates
Histogram for MCvM estimates of mu
Density
8
6
4
2
0
−0.1
0.0
0.1
Estimates
Histogram for AD estimates of mu
Density
10
5
0
−0.1
0.0
0.1
Estimates
Histogram for MAD estimates of mu
Density
12
9
6
3
0
−0.1
0.0
0.1
Estimates

$Figure : 17$



Histogram for ML estimates of sigma

Histogram for MLS estimates of sigma

Histogram for CvM estimates of sigma

Histogram for MCvM estimates of sigma

Histogram for AD estimates of sigma

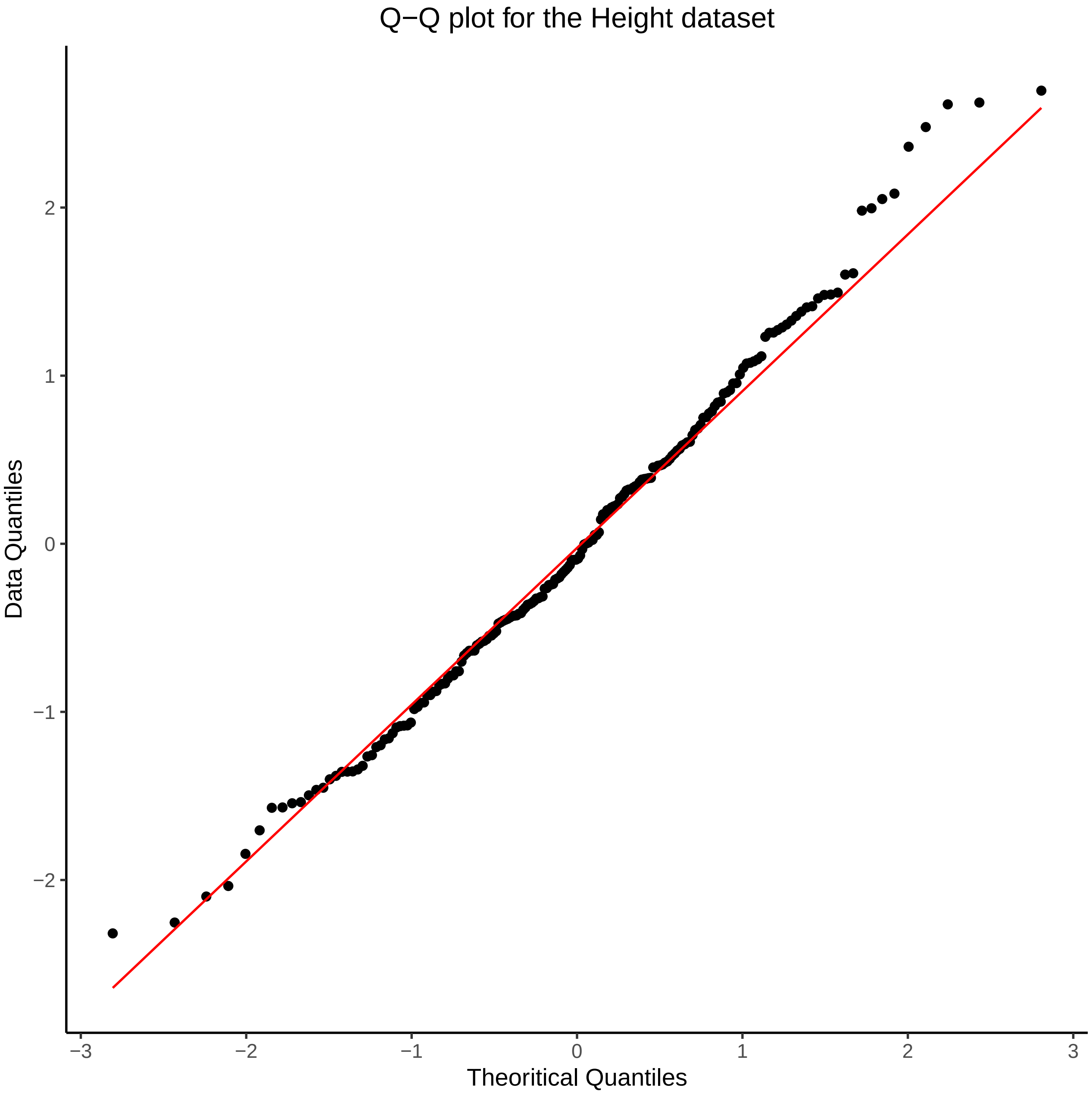
Histogram for MAD estimates of sigma

$Figure$: 18

NOTE : It can be observed that as the sample size increases, the MCvM and MAD estimators of the location parameter μ and scale parameter σ are performing very close to their counterparts CvM and AD estimators.

5 Real Data Analysis

Now, we will perform a real data analysis using “Height Dataset”. Now consider the following Q-Q plot of the data :



From the above Q-Q plot, we can observe that the quantiles of the standardized data perfectly matches with the quantiles of the $N(0,1)$ distribution. So we conclude that the standardized data comes from the Normal population and so does the original data. So here we will proceed further with the estimators used in the simulation study for the Normal (0,1) distribution and check how the estimators perform on real dataset.

Now to check the performance of the different estimators for our dataset, we adopt the following methodology :

- We will standardize the original datapoints by subtracting sample mean and dividing them by sample standard deviation.
- Then we will order the standardized data in ascending order.
- Now from the ordered data, we will estimate the population parameters μ and σ using the different estimation methods used in our simulation study (here, population corresponds to the standardised version of the original data).
- By adjusting the obtained estimators using sample mean and s.d. suitably, we will finally obtain the estimates of the parameters μ and σ of the original population distribution. Let's say the estimated parameter vector be $\widehat{\Theta}$.
- Now we will use the following formula of RMSE to evaluate the performance of different estimation methods:-

$$RMSE = [\frac{1}{n}\sum_{i=1}^{n}(F(x_{(i)};\widehat{\Theta}) - \frac{1}{n+1})^2]^{1/2}$$

where n is the number of samples and $F(\cdot)$ is the cumulative distribution function of Normal Distribution. The smaller value of the RMSE implies better fitting. Parameter estimates and RMSE values are tabulated in Table 10.

```
     mu.estimate sigma.estimate       RMSE
ML      68.11375       1.806430 0.01629942
MLS     68.05161       1.846098 0.01240468
CvM     68.05232       1.818682 0.01239811
MCvM    68.05159       1.832241 0.01233072
AD      68.07344       1.818512 0.01289939
MAD     68.07336       1.827505 0.01282035
```

From the above table we can observe the following :

- In terms of RMSE, all the estimators are performing well.
- To be more specific , MCvM outperforms all the other estimators.
- Also, it should be noted that, the modified estimators perform better than their original counterparts for this real data application.

6 Conclusions

- From the simulation study, it was shown that, as sample size increases, the MCvM and MAD estimators performed more or less same as their original counterparts. Though there were a little bit variations, in most of the cases, the Def values for all the estimators varied within a small range.

- For $N(0,1)$ case, if we consider the overall performance, MLE outperformed all the other estrimators in terms of bias, MSE and Def values of the estimators of μ and σ for all the three sample sizes. Also, MLSE performed worst in all the three sample sizes. Also there were no significant betterment in performance for the modified versions of CvM and AD estimators in terms of bias, MSE and Def.

- The same scenerio repeats in case of $EV(0,1)$ distribution and $Weibull(0,1,1.2)$ distribution.

- But the AD and MAD estimators also provide very good results for Weibull distribution.

- In real data application, it was observed that the MCvM performed best among all the estimators.Though the other estimators were also performing very well.

- In the original paper which we are using as our reference, there it was shown that in terms of average CPU times and number of iterations criteria, the Modified estimators perform better than the original ones. This part we could not cover in this project. But we can easily say that the modified estimators provide better results in some cases, sometimes they are as efficient as their original counterparts. Moreover computational cost is much less for them as shown in the original paper.

7 Contributions

- Paper finding : Anis Pakrashi, Krishnendu Paul, Souraj Mazumdar, Rahul Ghosh Dastidar

- Theory and derivation : Souraj Mazumdar, Krishnendu Paul

- Code : Anis Pakrashi

- Project Report : Rahul Ghosh Dastidar, Souraj Mazumdar

- Presentation : Krishnendu Paul, Rahul Ghosh Dastidar

8 References

1. Modified minimum distance estimators: definition, properties and applications by Talha Arslan1, Sukru Acitas2 , Birdal Senoglu3

2. Acitas S, Arslan T (2020) A comparison of different estimation methods for the parameters of the Weibull Lindley distribution. Eskisehir Teknik Üniv Bilim ve Teknol Dergisi B Teorik Bilimler 8(1):19–33

3. Basu A, Shioya H, Park C (2011) Statistical inference-the minimum distance approach. CRC Press, New York

4. Boos DD (1982) Minimum Anderson-Darling estimation. Commun Stat Theory Methods 11(24):2747– 2774