INDIAN INSTITUTE OF TECHNOLOGY,
MADRAS

CONCEPTS IN STATISTICAL LEARNING
PROGRAMMING ASSIGNMENT 1

CS6464

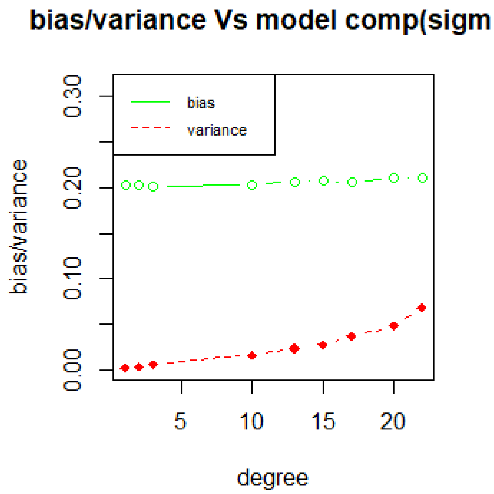
Linear Models

Author	*Roll Number*
DODIYA KEVAL	CS19M023

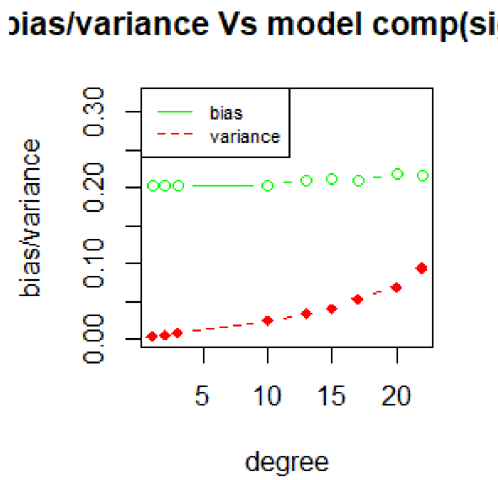
February 25, 2020

1 Task 1

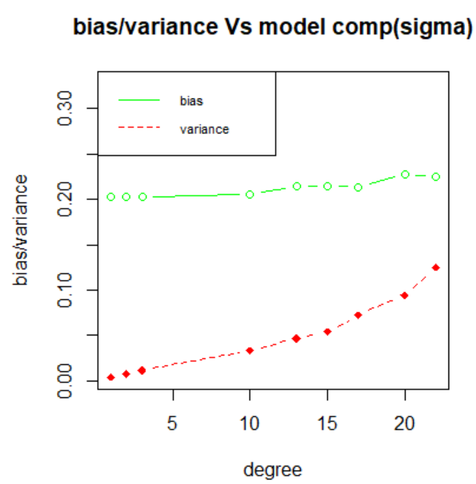
1.1 Function 1



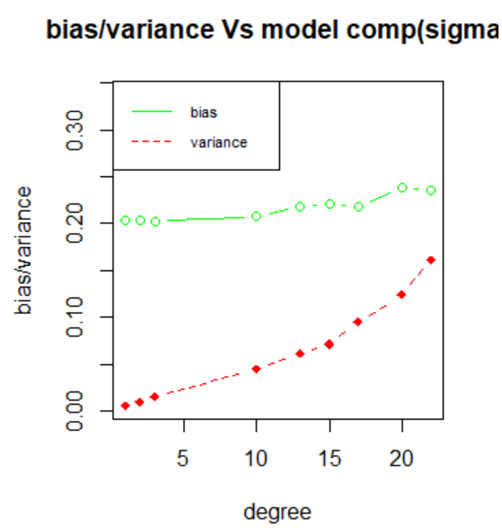
(a) $\sigma = 0.4$



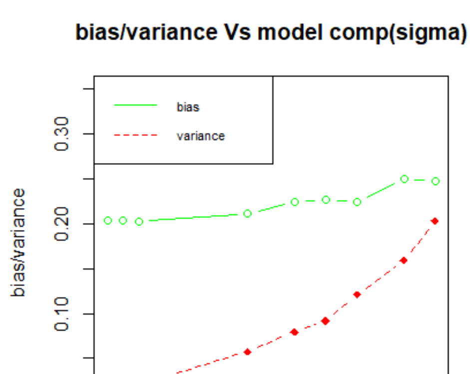
(b) $\sigma = 0.5$

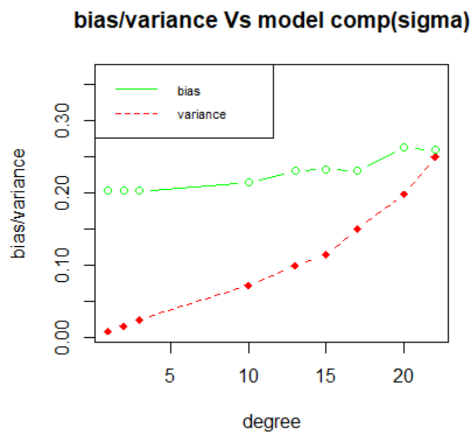


(c) $\sigma = 0.6$

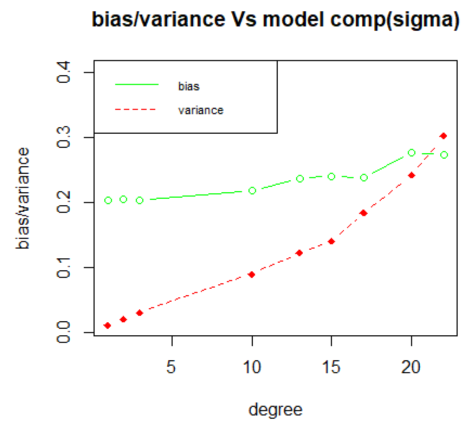


(d) $\sigma = 0.7$

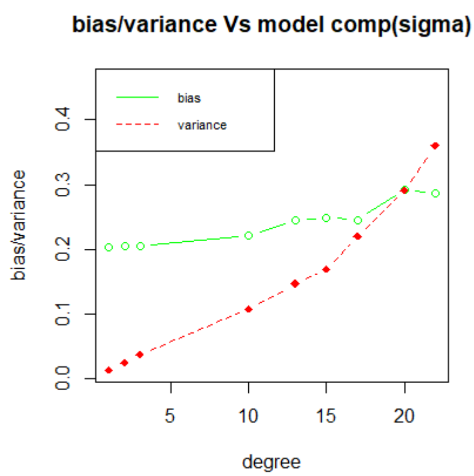




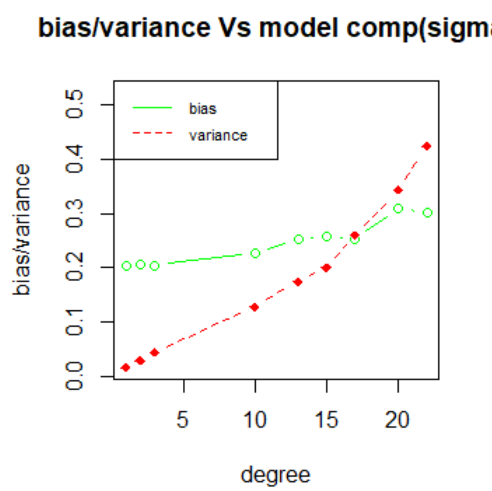
(a) $\sigma = 0.9$



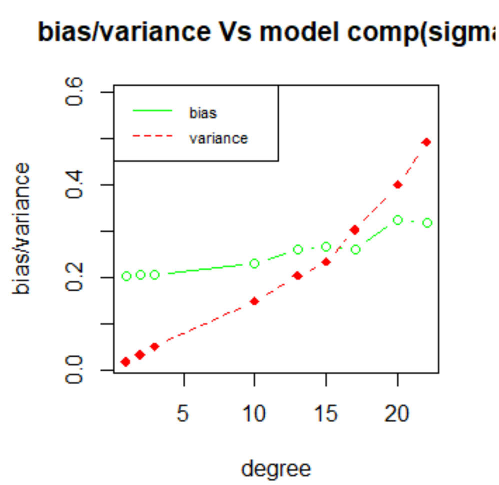
(b) $\sigma = 1$



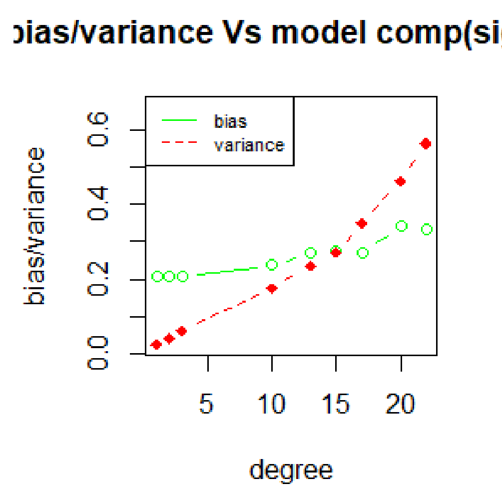
(c) $\sigma = 1.1$



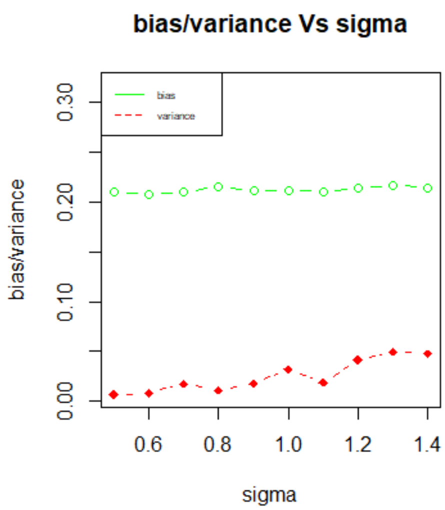
(d) $\sigma = 1.2$



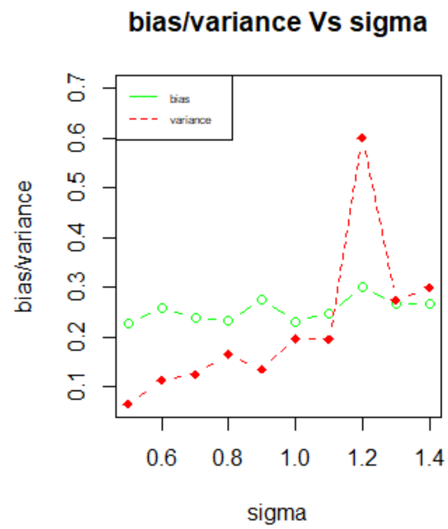
(e) $\sigma = 1.3$



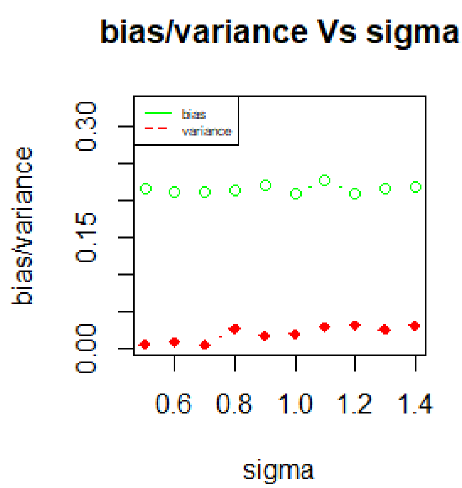
(f) $\sigma = 1.4$



(a) $d = 2$



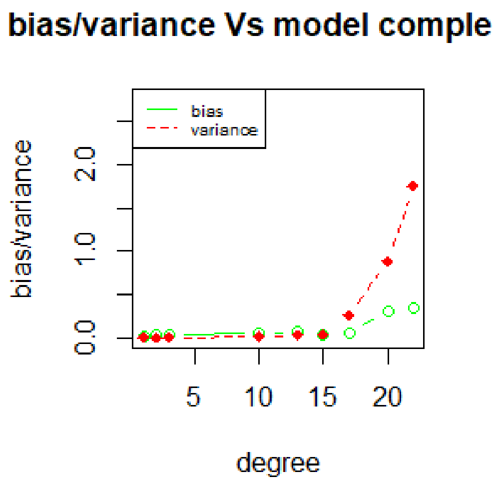
(b) $d = 13$



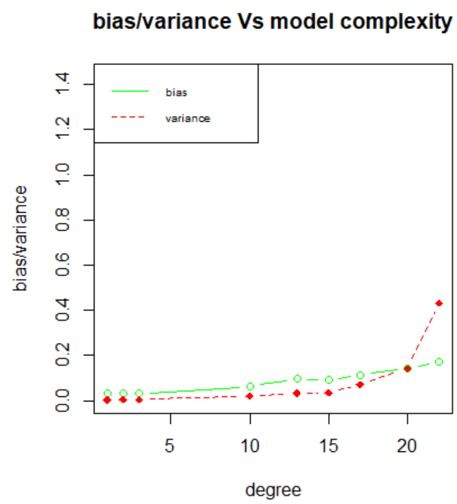
(c) $d = 23$

Figure 3: Error vs Sigma for d = {2,13,23}

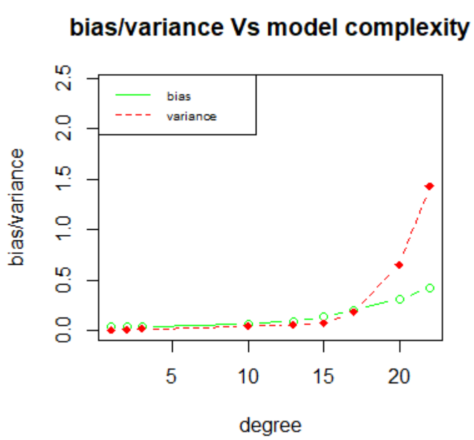
1.2 Function 2



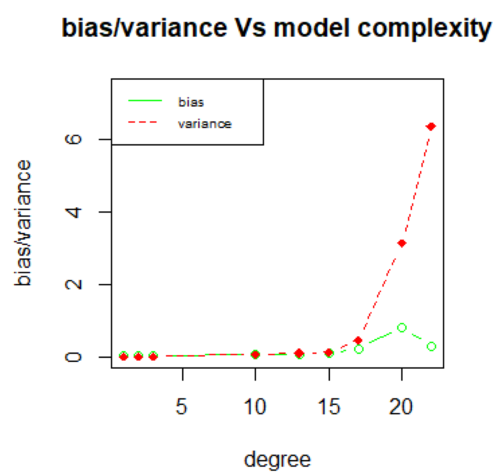
(a) $\sigma = 0.4$



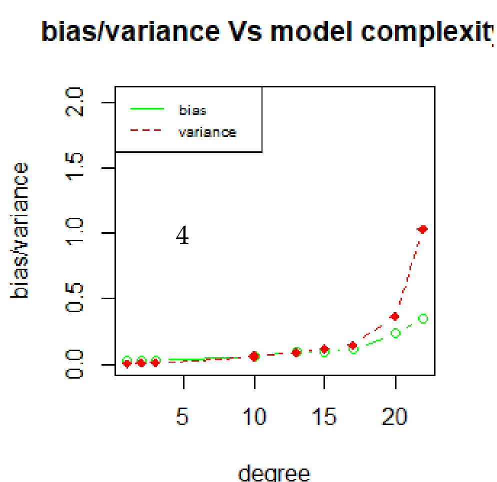
(b) $\sigma = 0.5$

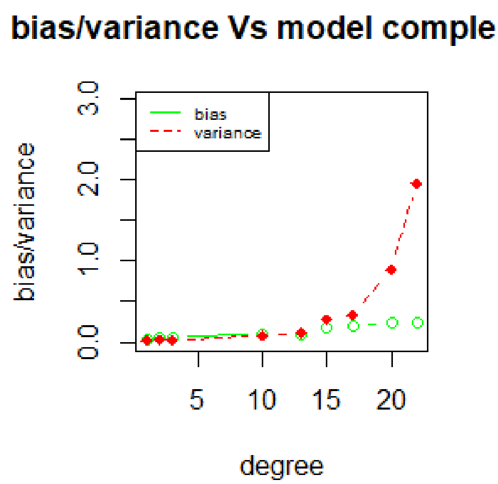


(c) $\sigma = 0.6$

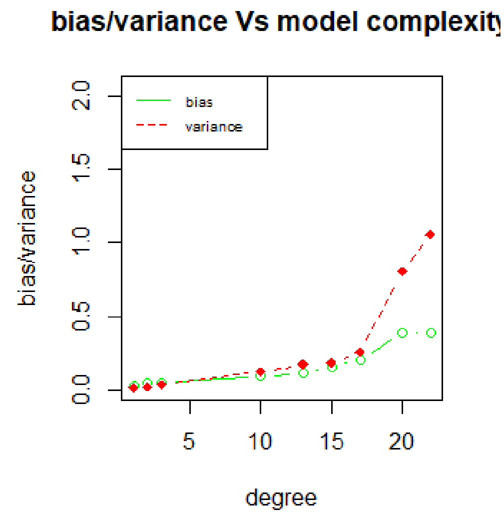


(d) $\sigma = 0.7$

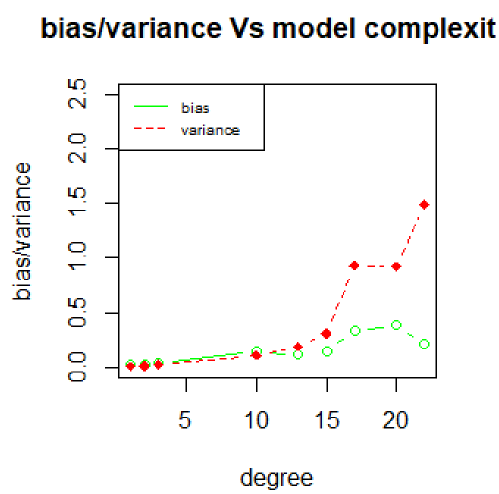




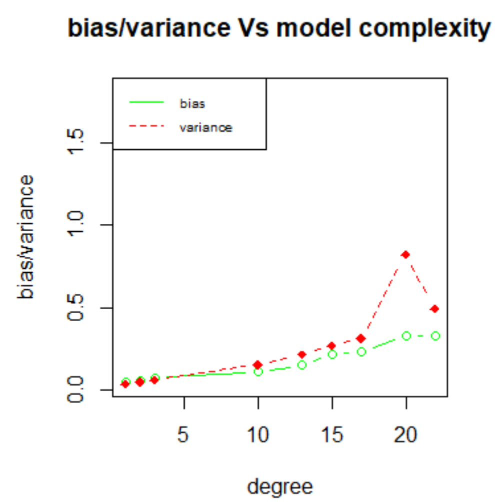
(a) $\sigma = 0.9$



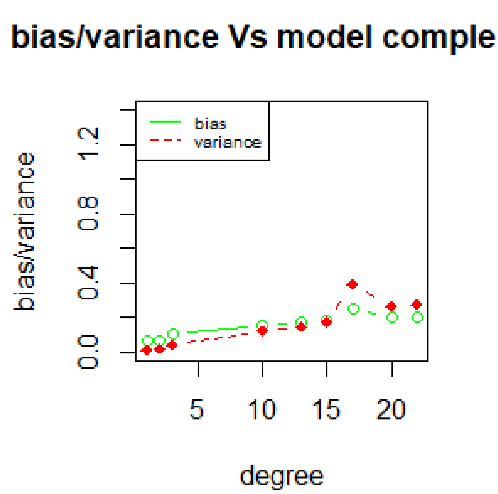
(b) $\sigma = 1$

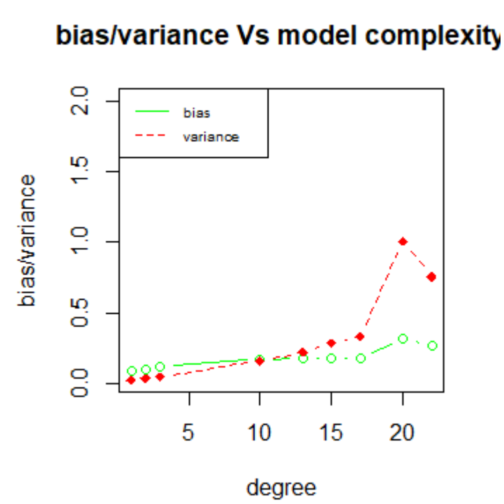


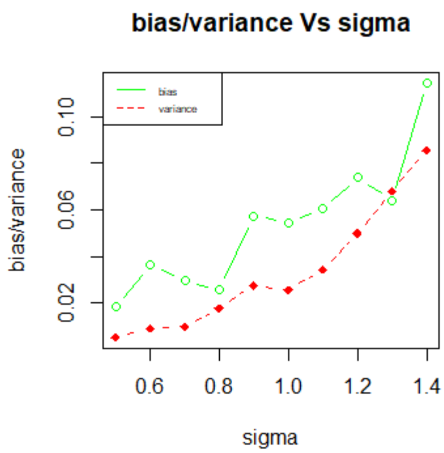
(c) $\sigma = 1.1$



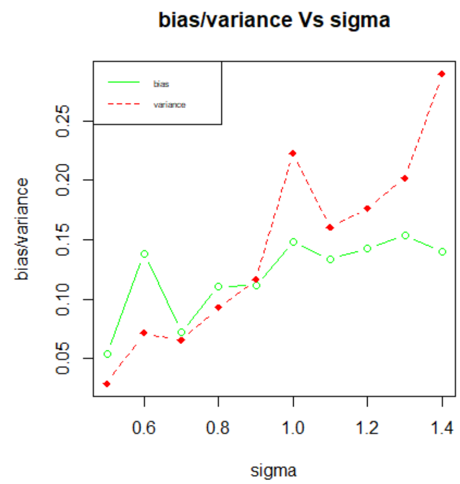
(d) $\sigma = 1.2$



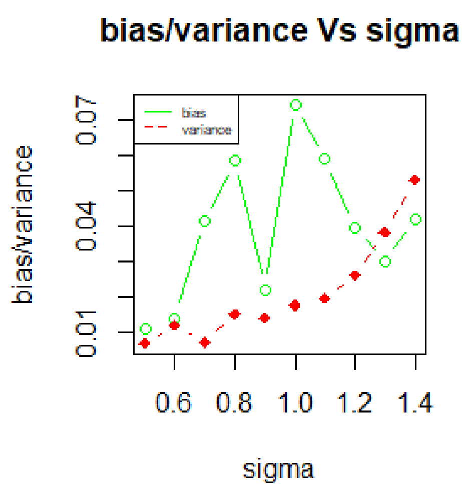




(a) $d = 2$



(b) $d = 13$



(c) $d = 23$

Figure 6: Error vs Sigma for d = {2,13,23}

function 1,2 given we needed to observe variance and bias behaviour.

- In both the functions we found out that when model complexity increases bias decreases and variance also increases and also MSE error also gets increases. sometimes it also gives different behaviour because of random data prediction and training.

- If we fix the model and varies sigma .variance gets increases heavily but bias behaviour is not conclusive but we can say for sure that variance always will increase. reason behind this behaviour is when sigma is high implies model's data gets more spread that indicate more variance.
- for second function i observed data is linear so at degree =2 only model gets overfitt hence variance will be more at more value of d(d =23).

2 Task 2

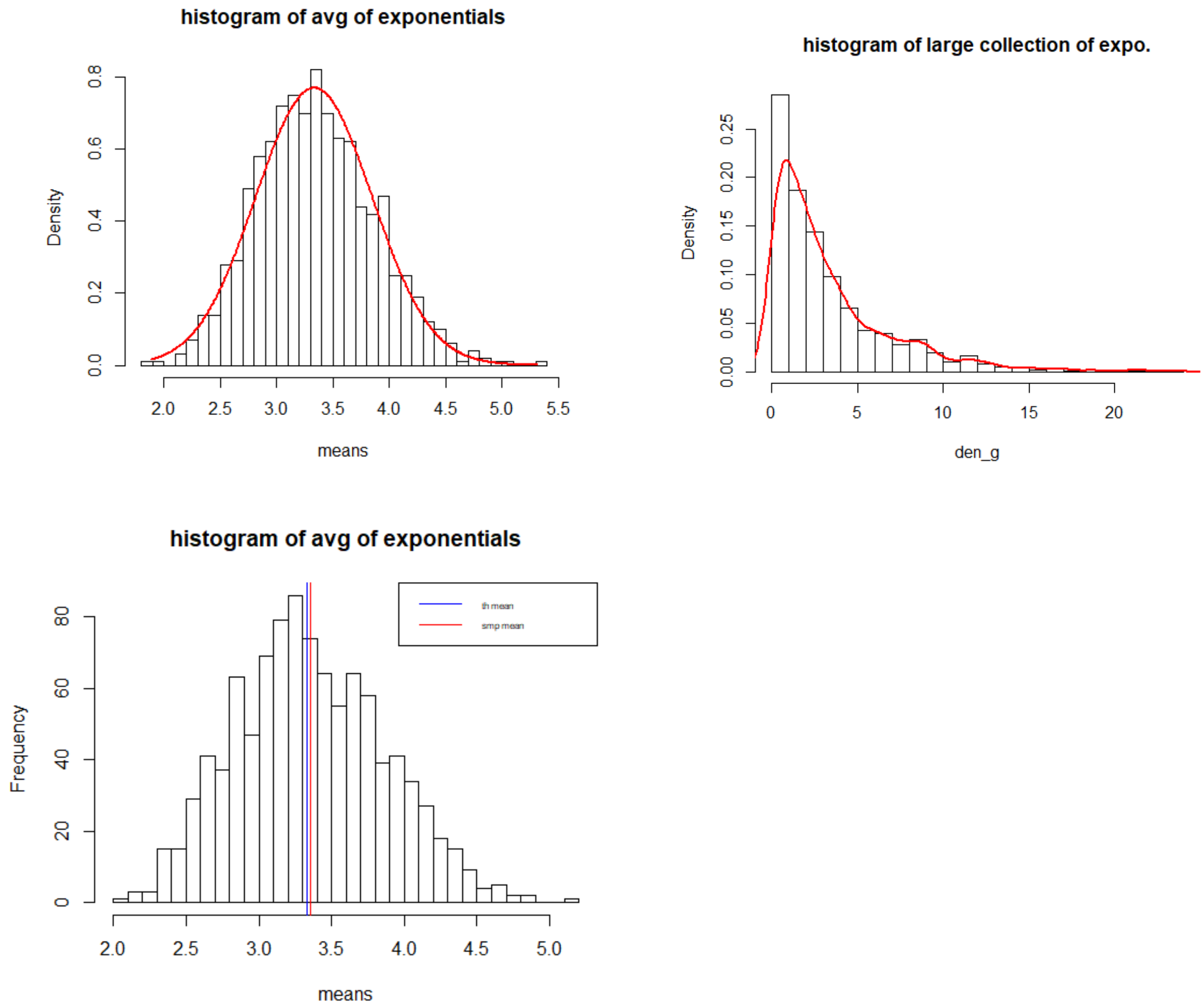


Figure 7: histogram of avg of exponential, histogram of large collection of exponentials(1000) ,comparison of theoretical mean and sample mean

- As we knows that theoretical of exponential distribution is $\frac{1}{\lambda}$, that is $1/0.3 = 3.333$. and the mean found by simulation is 3.35 and formulae given in code. and as we can see both are nearly same

- The theoretical variance of the exponential distribution is $\frac{\frac{1}{\lambda^2}}{n}$. here n is total samples. so theoretical variance will be 0.2777 and sample variance i got is 0.285 so both are close to each other

- first histogram describes large collection of 40 exponential(follows central limit theorem) . and second histogram describes large collection of exponentials

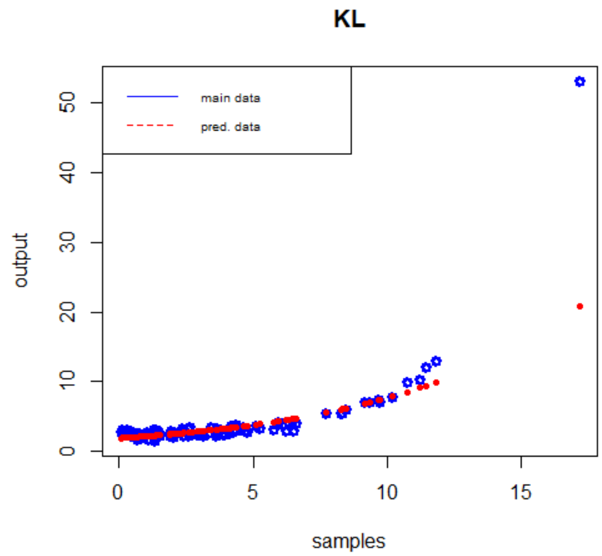
- so we can conclude that by the Central Limit Theorem, when we take large avg of exponentials. it gets properly normalized. so final distribution nearly obey normal distribution. The distribution would be even closer to a standard normal distribution if the sample size increased

2.1 parameter calculation

fisher information:

- Fisher information of exponential distribution can be find out by $\frac{n}{\lambda^2}$ so it'll be $(mean)^2 * n$ so Fisher information for sample mean we got as 450 and for theoretical mean it'll be 444.44 near to same ,

KL divergence:



figure(a): variance in plotted exponential function

- KL divergence uses to find out how much both the model differs its one kind of entropy. suppose we have two models(distribution) to compare P ,Q then KL divergence will be..

$$\sum_{1}^{n} P(X_i)(log(\frac{P(X_i)}{Q(X_i)})) \tag{1}$$

- the KL value I got for fitted data is vary small like $-3.61 * 10^{-12}$.this is because data is totally fitted and i got large KL divergence value when i added variance(sample from normal distribution with mean = 2, sd = 0.4) in exponential function and it is like 178(means prediction is less accurate or high variance) that thing discribes in figure (a)
- reference for KL divergence is https://www.countbayesie.com/blog/2017/5/9/kullback-leibler-divergence-explained
- reference for fisher information https://en.wikipedia.org/wiki/$Fisher_information$
- reference task2 https://rpubs.com/hugoandrade/331744