

Problem2

2022-11-14

Problem 2 : Simulation Study to Understand Sampling Distribution

Part A Suppose $X_1, X_2, \dots, X_n \stackrel{iid}{\sim} \text{Gamma}(\alpha, \sigma)$, with pdf as

$$f(x|\alpha, \sigma) = \frac{1}{\sigma^\alpha \Gamma(\alpha)} e^{-x/\sigma} x^{\alpha-1}, \quad 0 < x < \infty,$$

The mean and variance are $E(X) = \alpha\sigma$ and $\text{Var}(X) = \alpha\sigma^2$. Note that **shape** = α and **scale** = σ . 1. Write a function in R which will compute the MLE of $\theta = \log(\alpha)$ using **optim** function in R. You can name it MyMLE

```
My_MLE <- function(x){
  ll_gamma1 <- function(p){
    ll_gamma(p,x)
  }
  z <- optim(par=c(1,1),ll_gamma1)
  log(z$par[1])
}
```

2. Choose n=20, and alpha=1.5 and sigma=2.2

- (i) Simulate $\{X_1, X_2, \dots, X_n\}$ from **rgamma**(n=20,shape=1.5,scale=2.2)
- (ii) Apply the MyMLE to estimate θ and append the value in a vector
- (iii) Repeat the step (i) and (ii) 1000 times
- (iv) Draw histogram of the estimated MLEs of θ .
- (v) Draw a vertical line using **abline** function at the true value of θ .
- (vi) Use **quantile** function on estimated θ 's to find the 2.5 and 97.5-percentile points.

```
library(stats)
vector = c()
gamma_simulation_MLE <- function(N,sim_size,alpha,sigma){
  for (j in 1:sim_size)
  {
    x<-rgamma(N,shape=alpha,scale=sigma)

    ll_gamma <- function(p,x){
      alpha<-p[1]
      sigma<-p[2]
      N*alpha*log(sigma)+N*log(gamma(alpha))-(alpha-1)*sum(log(x))+sum(x)/sigma
    }

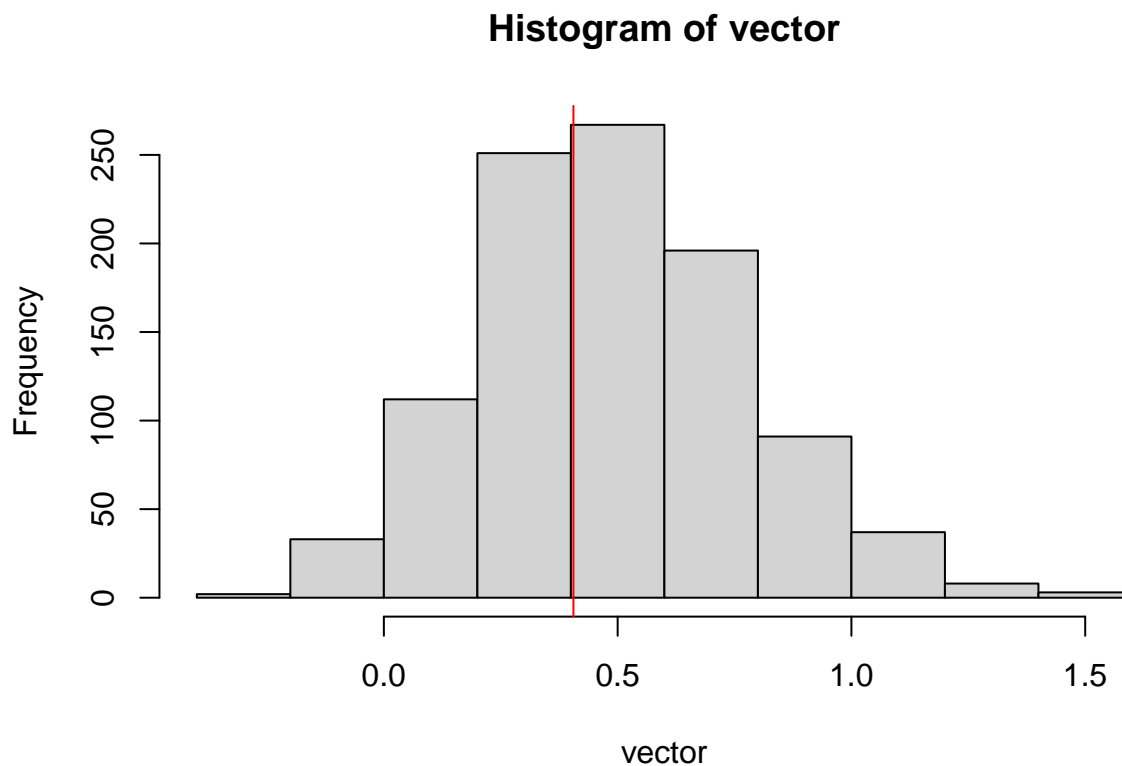
    My_MLE <- function(x){
      ll_gamma1 <- function(p){
        ll_gamma(p,x)
      }

```

```

    }
    z <- optim(par=c(1,1),ll_gamma1)
    log(z$par[1])
  }
  vector <- c(vector, My_MLE(x))
}
hist(vector)
abline(v=log(alpha),col='red')
cat("The 2.5 and 97.5 percentile points are:")
quantile(vector,probs=c(0.025,0.975))
}
gamma_simulation_MLE(20,1000,1.5,2.2)

```

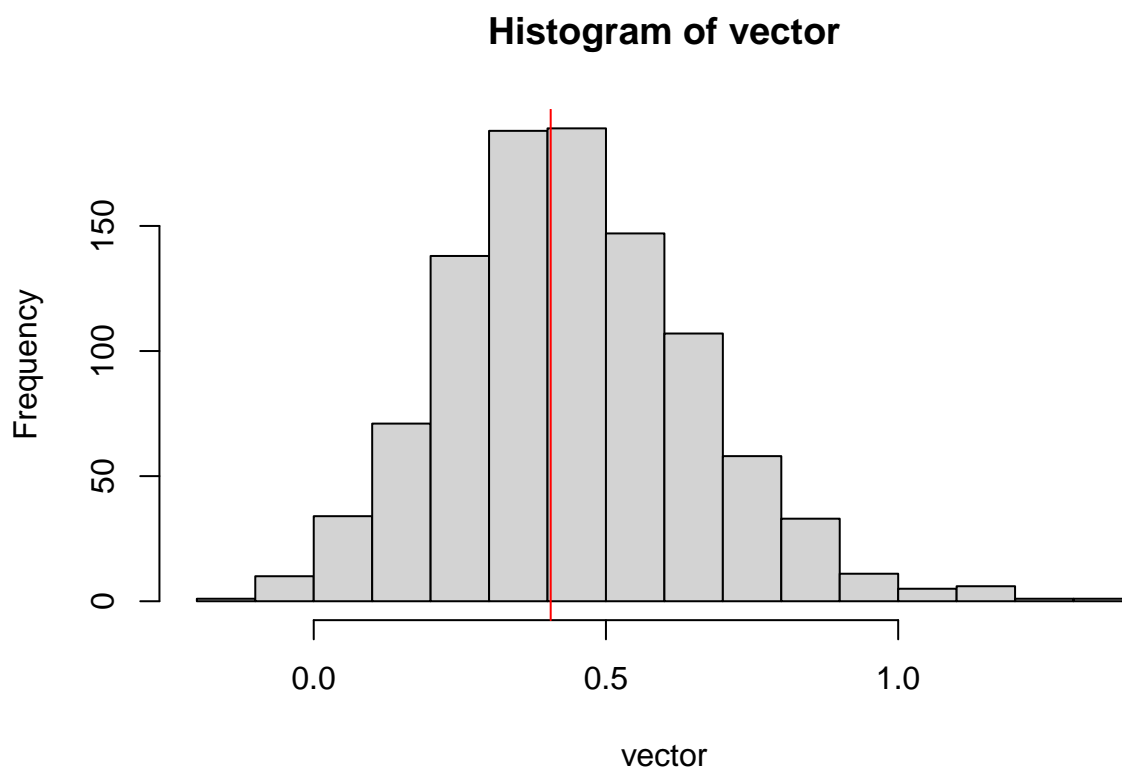


The 2.5 and 97.5 percentile points are:

```
##      2.5%      97.5%
## -0.039267  1.109965
```

3. Choose $n=40$, and $\alpha=1.5$ and repeat the (2).

```
gamma_simulation_MLE(40,1000,1.5,2.2)
```

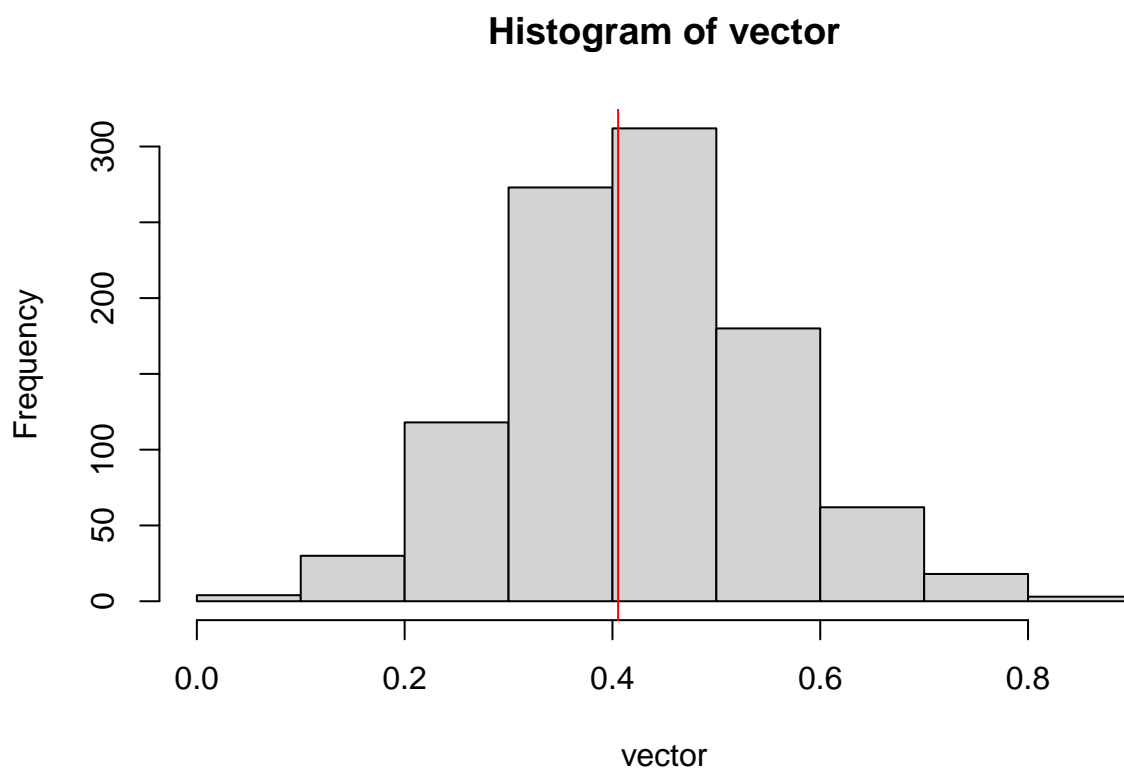


The 2.5 and 97.5 percentile points are:

```
##      2.5%      97.5%  
## 0.04333094 0.89140559
```

4. Choose $n=100$, and $\alpha=1.5$ and repeat the (2).

```
gamma_simulation_MLE(100,1000,1.5,2.2)
```



```
## The 2.5 and 97.5 percentile points are:
```

```
##      2.5%      97.5%  
## 0.1864060 0.6873139
```

5. Check if the gap between 2.5 and 97.5-percentile points are shrinking as sample size **n** is increasing?

Yes the gap between the 2.5 and 97.5 percentile points gets smaller as **n** increases.