Advanced Statistical Inference: Assignment 2

ankit.tewari@estudiant.upc.edu

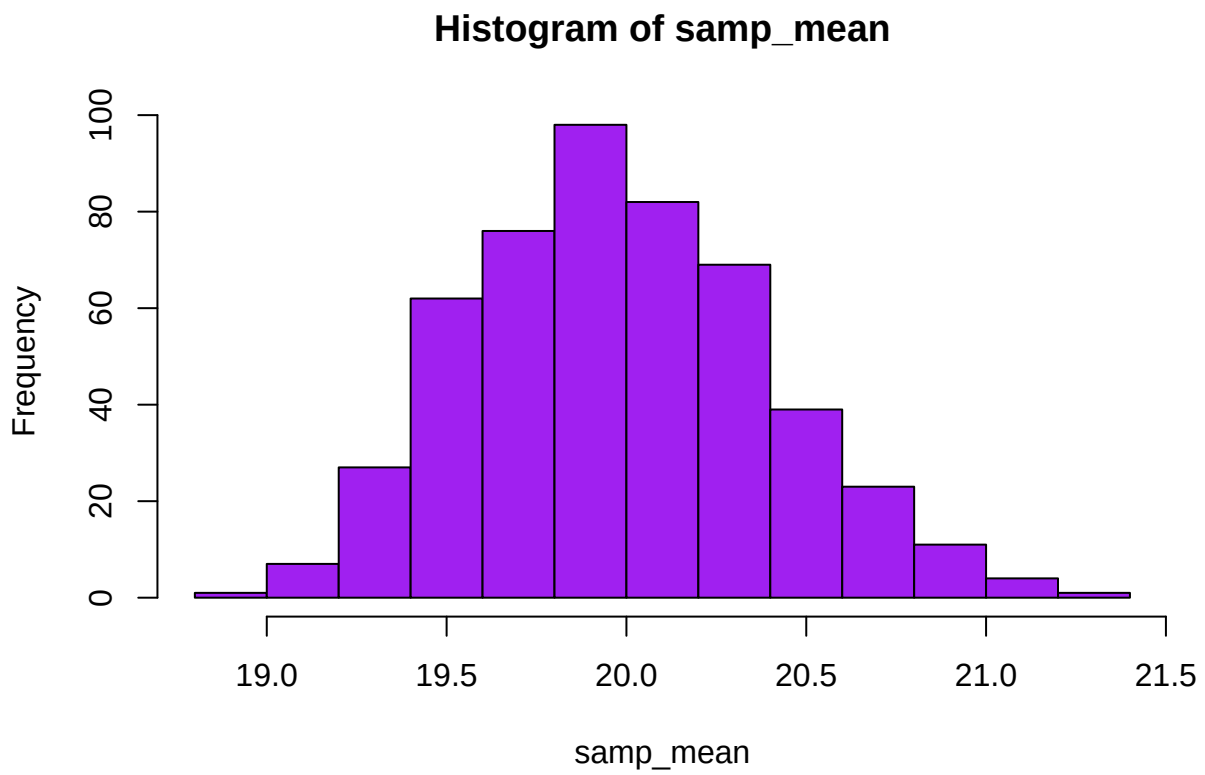
We're creating a population that has gamma distribution with paramters r=5 and lambda=1/4

```
dist<-rgamma(500, 5, 1/4)
mean_dist<-mean(dist)
```

Sampling Distribution of Means

```
samp_mean<-rgamma(500, 500*5, 500*(1/4))
samp_mu<-mean(samp_mean)
samp_sd<-sd(samp_mean)

hist(samp_mean, col = "purple")
```

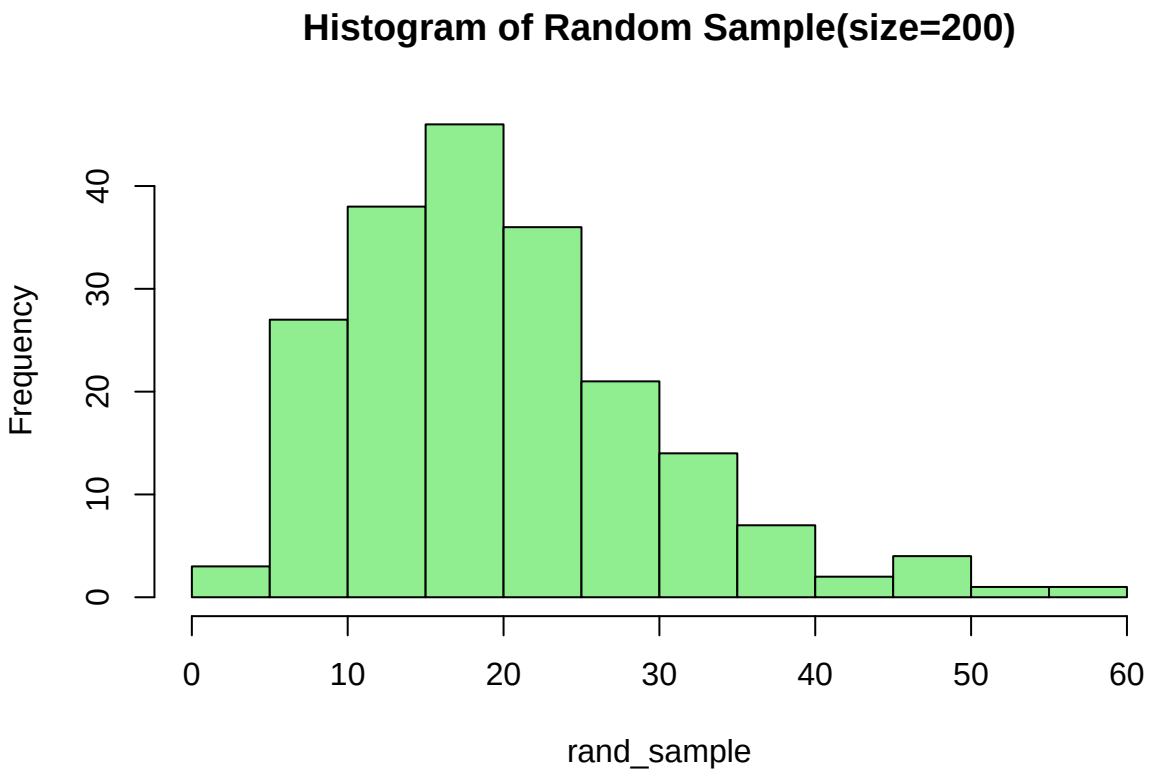


Creating a random sample of size 200 from the population

```
rand_sample<-sample(dist, 200, replace = F)

mean_rand_sample<-mean(rand_sample)
sd_rand_sample<-sd(rand_sample)

hist(rand_sample, col = "lightgreen", main = "Histogram of Random Sample(size=200)")
```

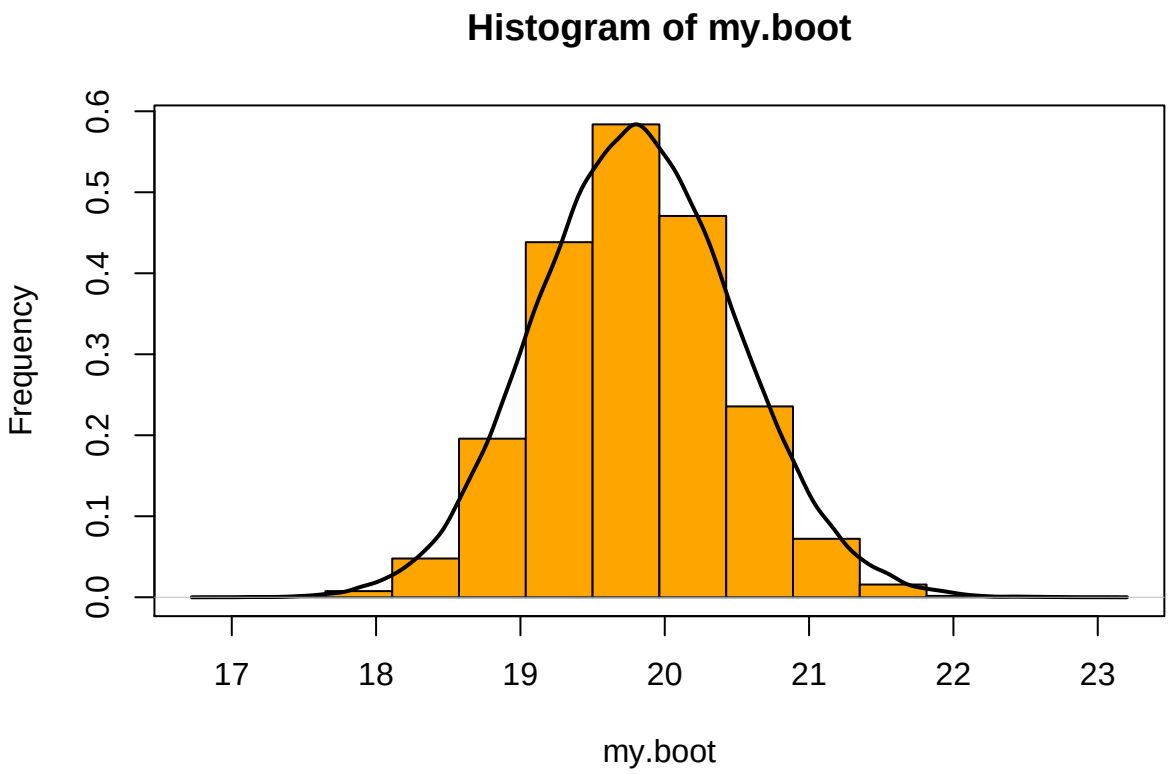


Bootstrap Distribution

```
N<-10^5
my.boot<-numeric(N)

for(i in 1:N){
  x<-sample(rand_sample, 200, replace = T)
  my.boot[i]<-mean(x)
}
mean_boot<-mean(my.boot)
sd_boot<-sd(my.boot)

hist(my.boot, col = "Orange", axes = F)
par(new=T)
plot(density(my.boot), main = "", xlab = "", ylab = "", lwd=2)
```



Comparison table:

```
r1<-c("Population", mean_dist, sd(dist))
r2<-c("Sampling Distribution of Mean", mean_value=samp_mu, sd_value=samp_sd)
r3<-c("Sample (size=200)", mean_rand_sample, sd_rand_sample)
r4<-c("Bootstrap Distribution", mean_boot, sd_boot)

result<- as.data.frame(rbind(r1, r2, r3, r4))
names(result)<-c("", "Mean", "Standard Deviation")
rownames(result) <- c()

result
```

```
##                                        Mean Standard Deviation
## 1                    Population 19.5612560431298   9.06330289034017
## 2 Sampling Distribution of Mean 19.9742215223764  0.407164975857717
## 3             Sample (size=200) 19.8050822460605   9.83659236054303
## 4        Bootstrap Distribution 19.8043186787529  0.691518016877316
```

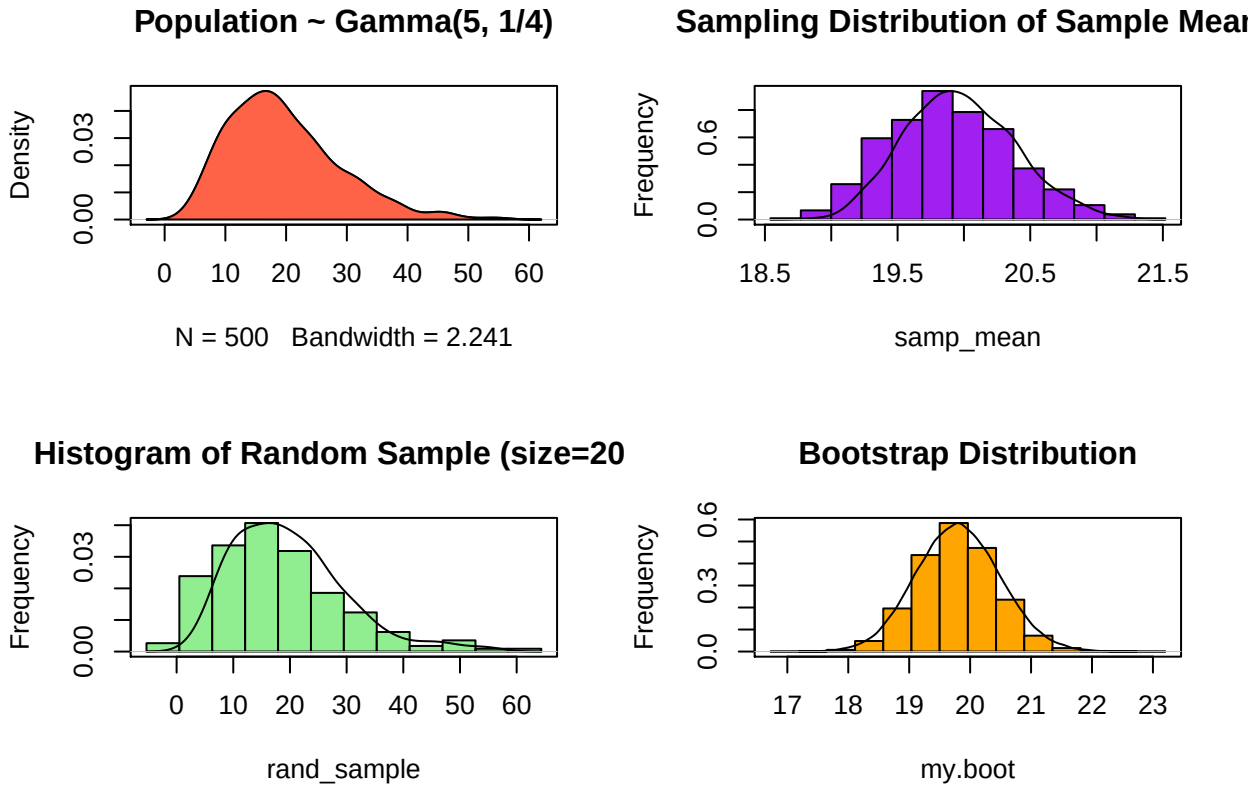
```
par(mfrow=c(2,2))

plot(density(dist), main = "Population ~ Gamma(5, 1/4)")
polygon(density(dist), col="tomato", border="black")

hist(samp_mean, col = "purple", main = "Sampling Distribution of Sample Mean", axes = F)
par(new=T)
plot(density(samp_mean), main = "", xlab = "", ylab = "", lwd=1)
```

```r
hist(rand_sample, col = "lightgreen", main = "Histogram of Random Sample (size=200)", axes = F)
par(new=T)
plot(density(rand_sample), main = "", xlab = "", ylab = "", lwd=1)

hist(my.boot, col = "Orange", axes = F, main = "Bootstrap Distribution")
par(new=T)
plot(density(my.boot), main = "", xlab = "", ylab = "", lwd=1)
```



for a sample size of 50,

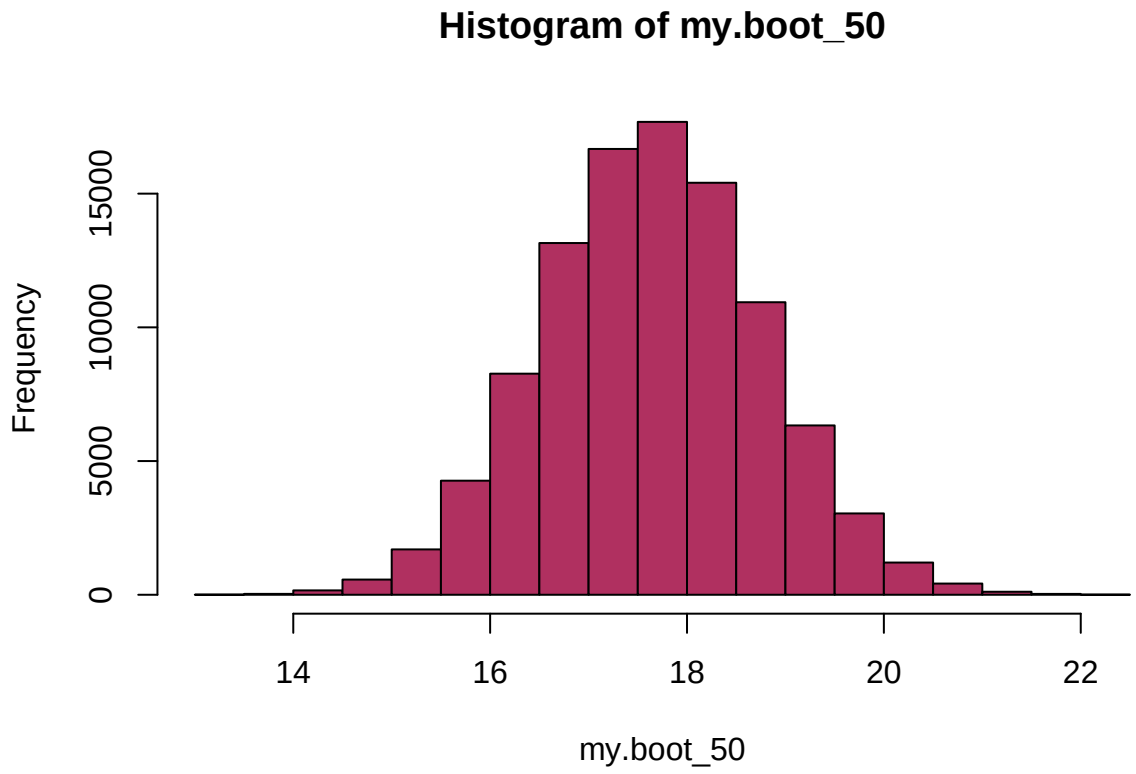
```r
rand_sample_50<-sample(dist, 50, replace = F)

mean_rand_sample_50<-mean(rand_sample_50)
sd_rand_sample_50<-sd(rand_sample_50)


N<-10^5
my.boot_50<-numeric(N)

for(i in 1:N){
  x<-sample(rand_sample_50, 50, replace = T)
  my.boot_50[i]<-mean(x)
}
mean_boot_50<-mean(my.boot_50)
sd_boot_50<-sd(my.boot_50)
```

```r
hist(my.boot_50, col = "maroon")
```



for sample size of 10

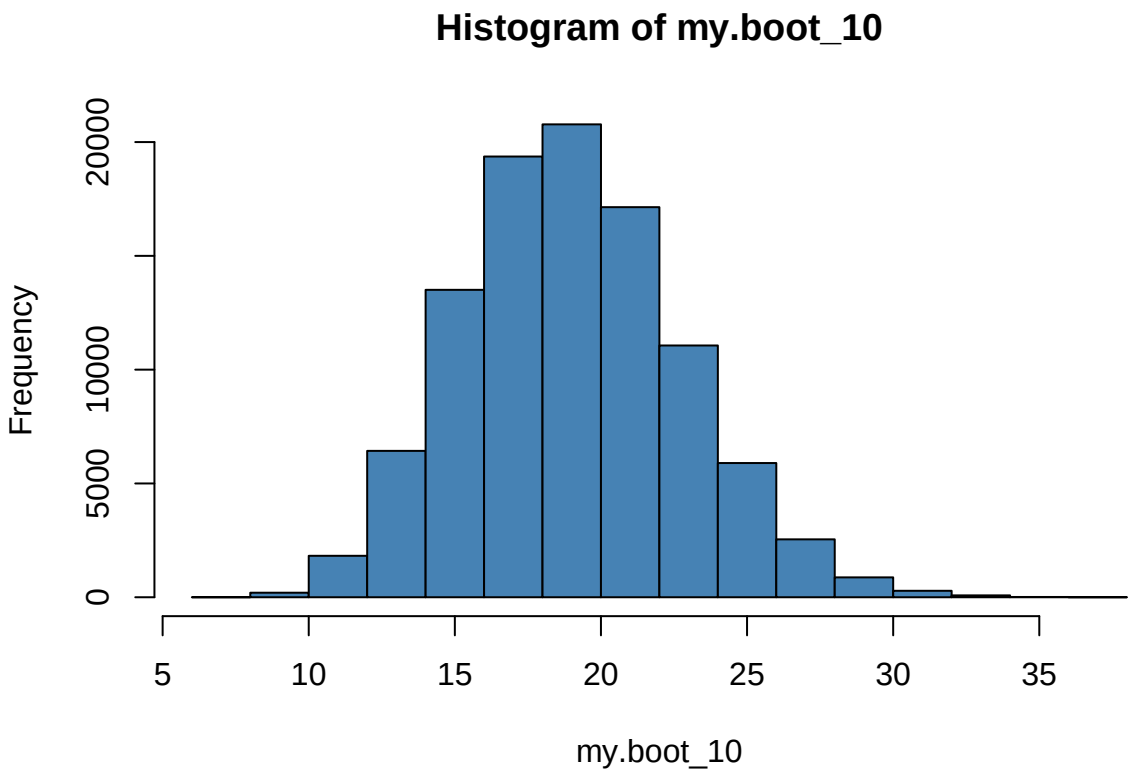
```r
rand_sample_10<-sample(dist, 10, replace = F)

mean_rand_sample_10<-mean(rand_sample_10)
sd_rand_sample_10<-sd(rand_sample_10)


N<-10^5
my.boot_10<-numeric(N)

for(i in 1:N){
  x<-sample(rand_sample_10, 10, replace = T)
  my.boot_10[i]<-mean(x)
}
mean_boot_10<-mean(my.boot_10)
sd_boot_10<-sd(my.boot_10)

hist(my.boot_10, col = "steel blue")
```


Histogram of my.boot_10
Frequency
0
5000
10000
20000
5
10
15
20
25
30
35
my.boot_10