CLT_HW Q5

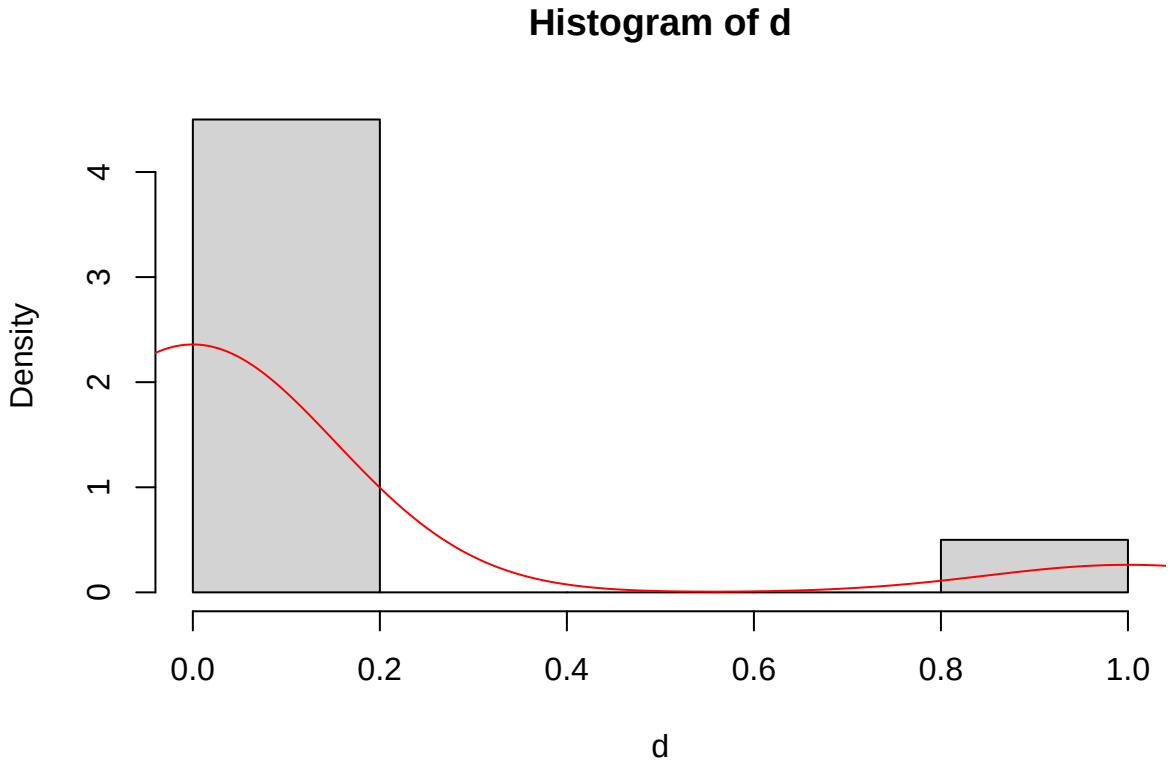
Aoyi Li

2021/3/5

Choose a skewed binomial distribution

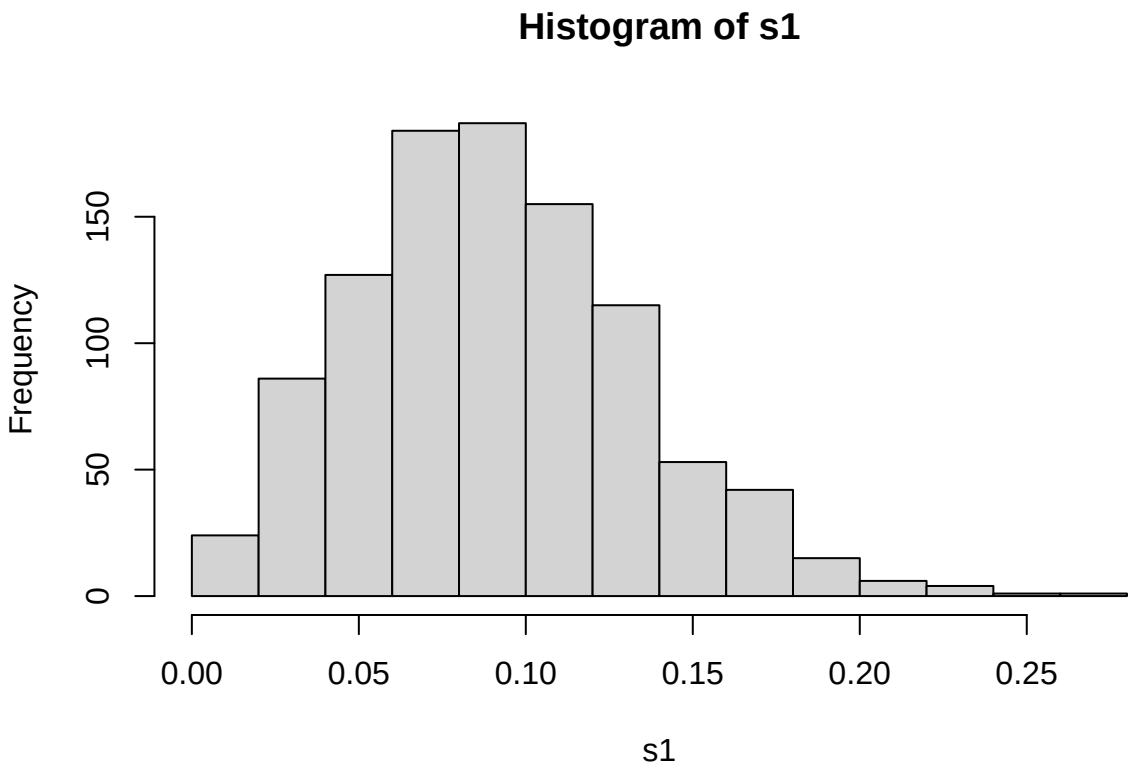
```
#The skewed binomial distribution
set.seed(677)
n<-20
p<-0.1
d<-rbinom(n,1,p)

#histogram of this distribution
hist(x = d, freq = FALSE)
lines(x = density(x = d), col = "red")
```

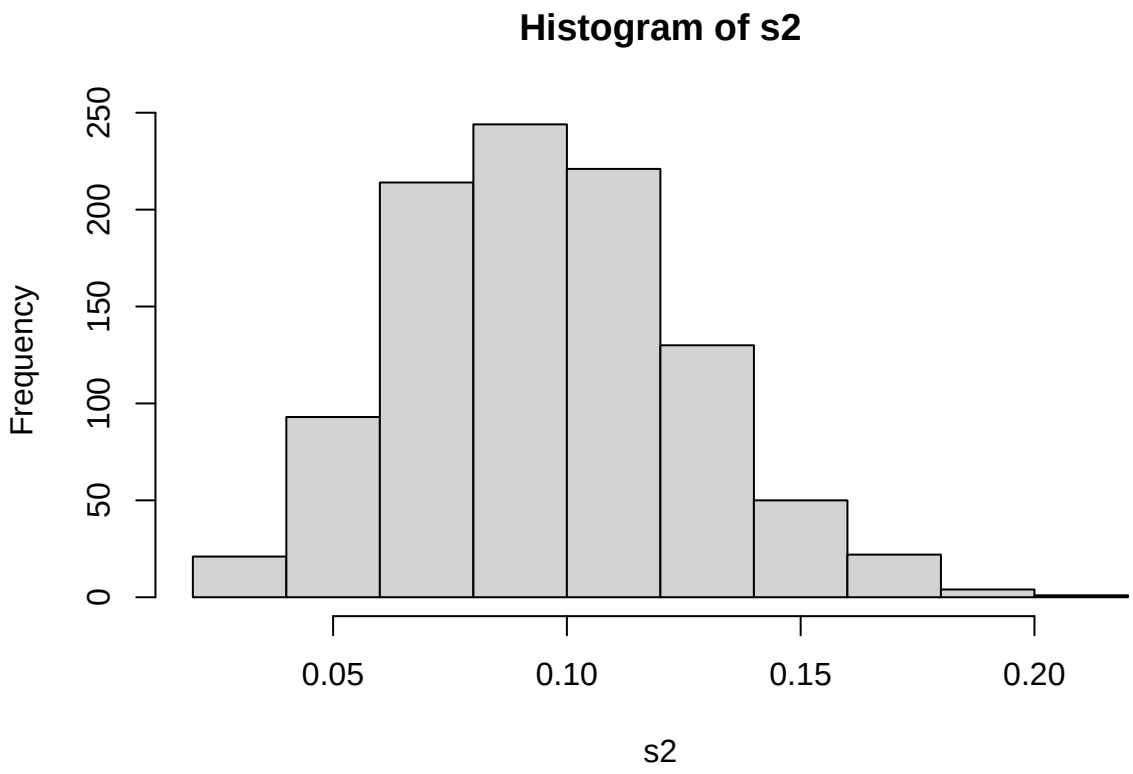


Normal approximation to the binomial distribution

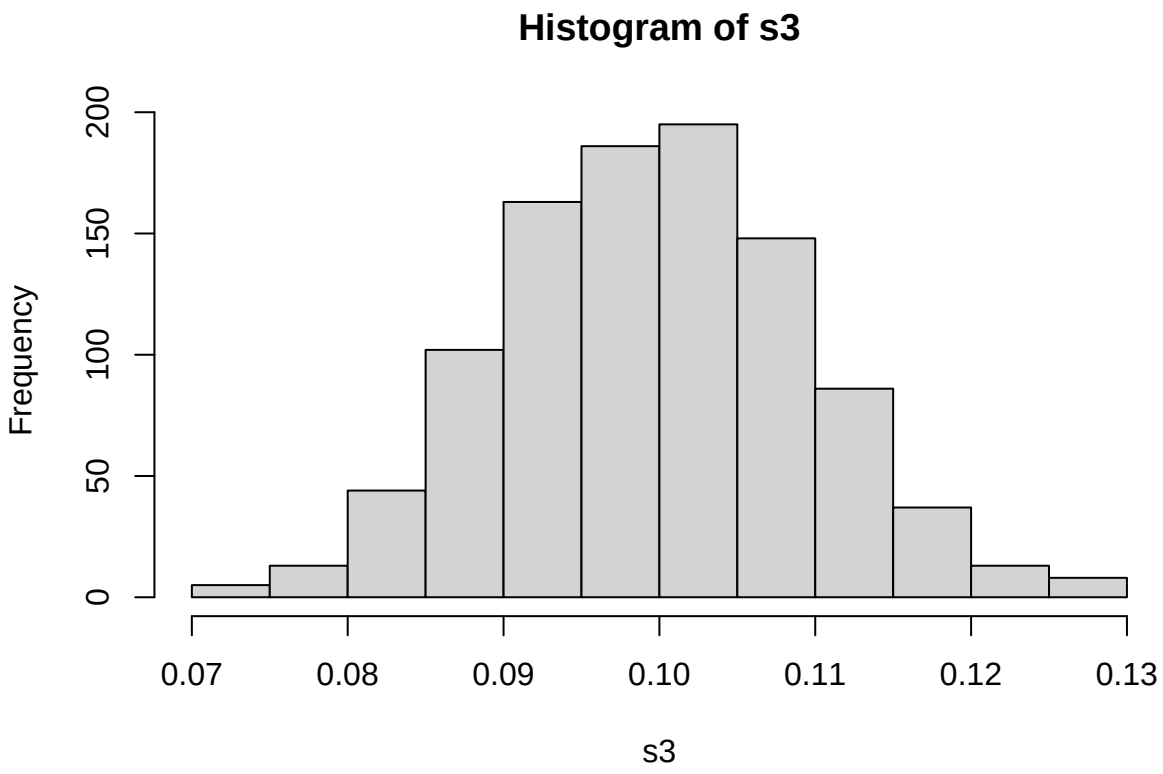
```r
#Visually check that when n groups up, the binomial distribution approximate to normal.
s1<-c()
for(i in 1:1000){
  n<-rbinom(50,1,p)
  s1[i]=mean(n)
}
hist(s1)
```



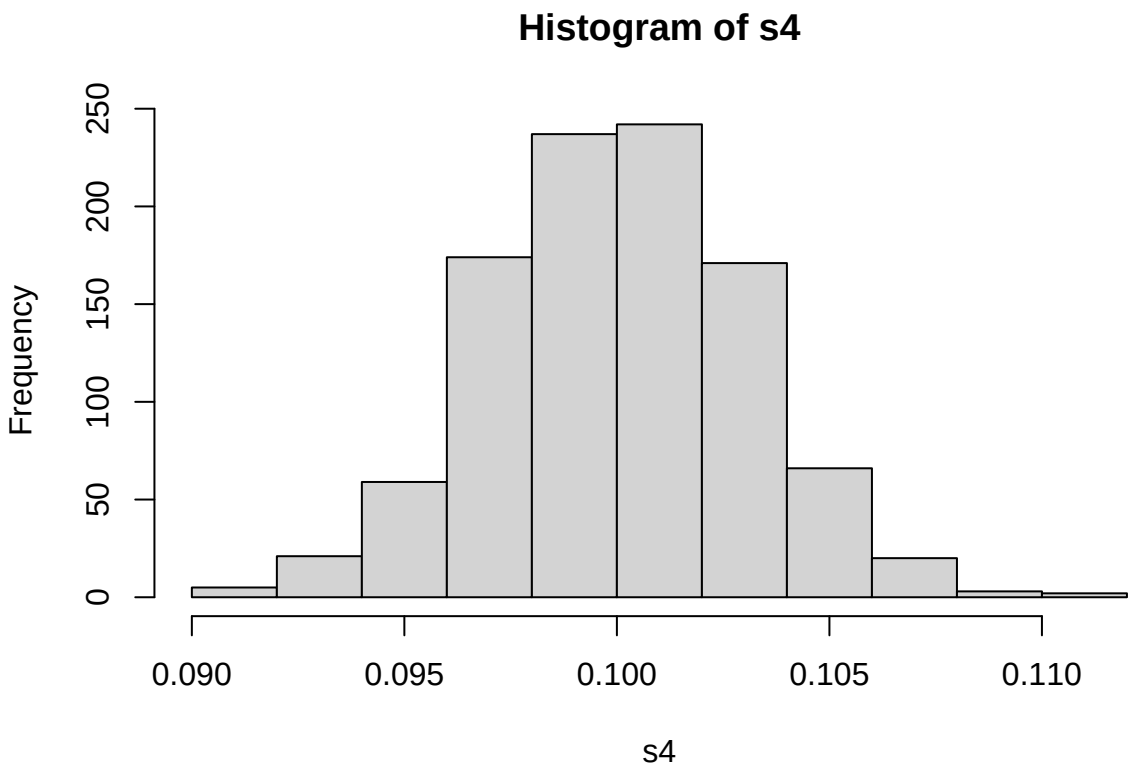

```r
s2<-c()
for(i in 1:1000){
  n1<-rbinom(100,1,p)
  s2[i]=mean(n1)
}
hist(s2)
```



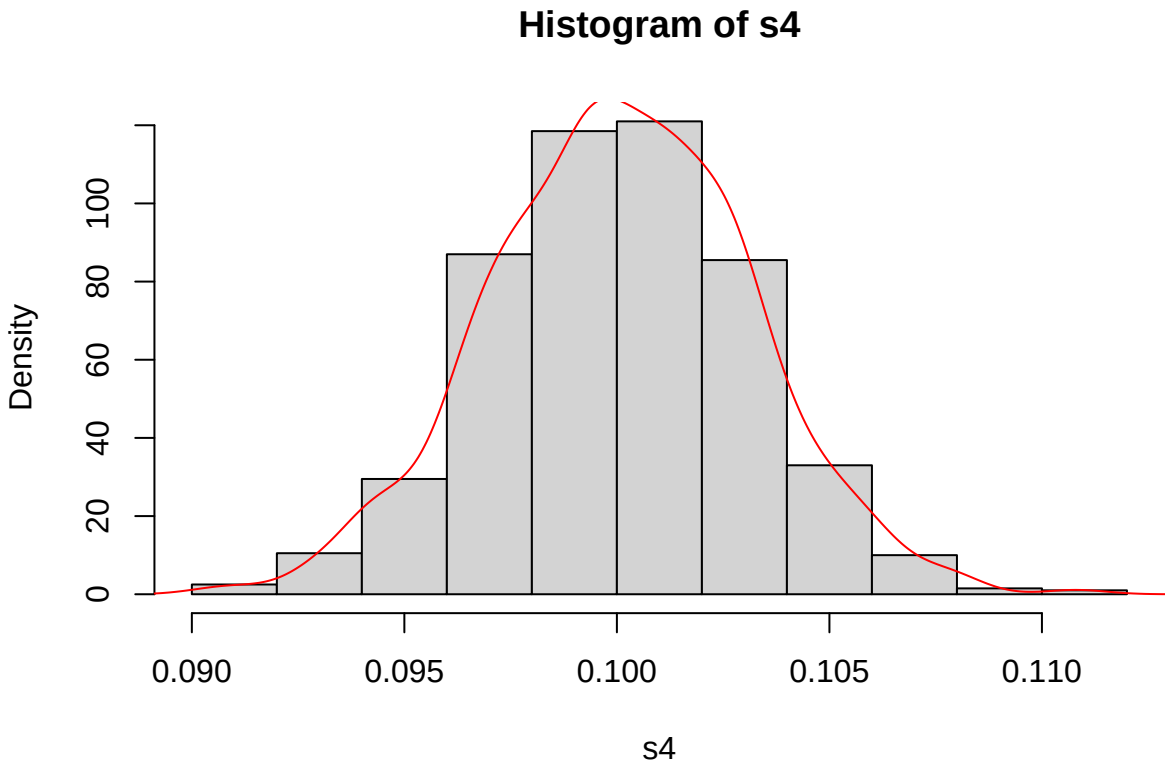

```
s3<-c()
for(i in 1:1000){
  n3<-rbinom(1000,1,p)
  s3[i]=mean(n3)
}
hist(s3)
```




```
s4<-c()
for(i in 1:1000){
  n4<-rbinom(10000,1,p)
  s4[i]=mean(n4)
}
hist(s4)
```



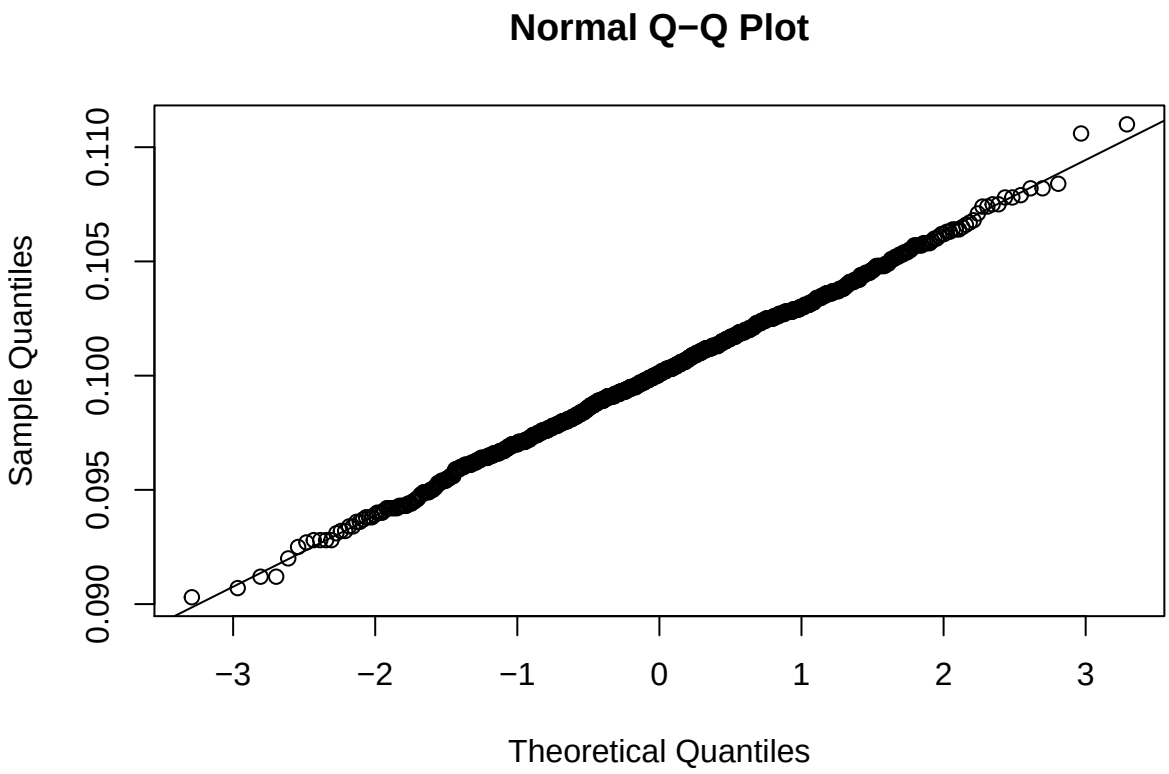

```
#normality check
hist(x = s4, freq = FALSE)
lines(x = density(x = s4), col = "red")
```




```
shapiro.test(s4)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  s4
## W = 0.99887, p-value = 0.7968
```

```
qqnorm(s4)
qqline(s4)
```



The histogram for the generated data shows a belled curve. The shapiro test for the distribution with n=10000 shows p-value greater than 0.05, which indicates the sample is normally distributed. The Q-Q Plot also shows a straight line showing the normality.