

STAT__5034 HW4

zhengzhi lin

2019.10.21

P1

a As n increases, the variance of sampling distribution of sample mean decreases.

c In this question, the mean of CLT normal distribution is the mean of beta distribution, which is controlled by α , the variance of beta distribution is related to α , and the variance of CLT normal distribution is the variance of beta times $1/n$, thus the variance of CLT normal is controlled by α and n . We can see from the plots, as n increases, the hist plot is more matched by the normal curve.

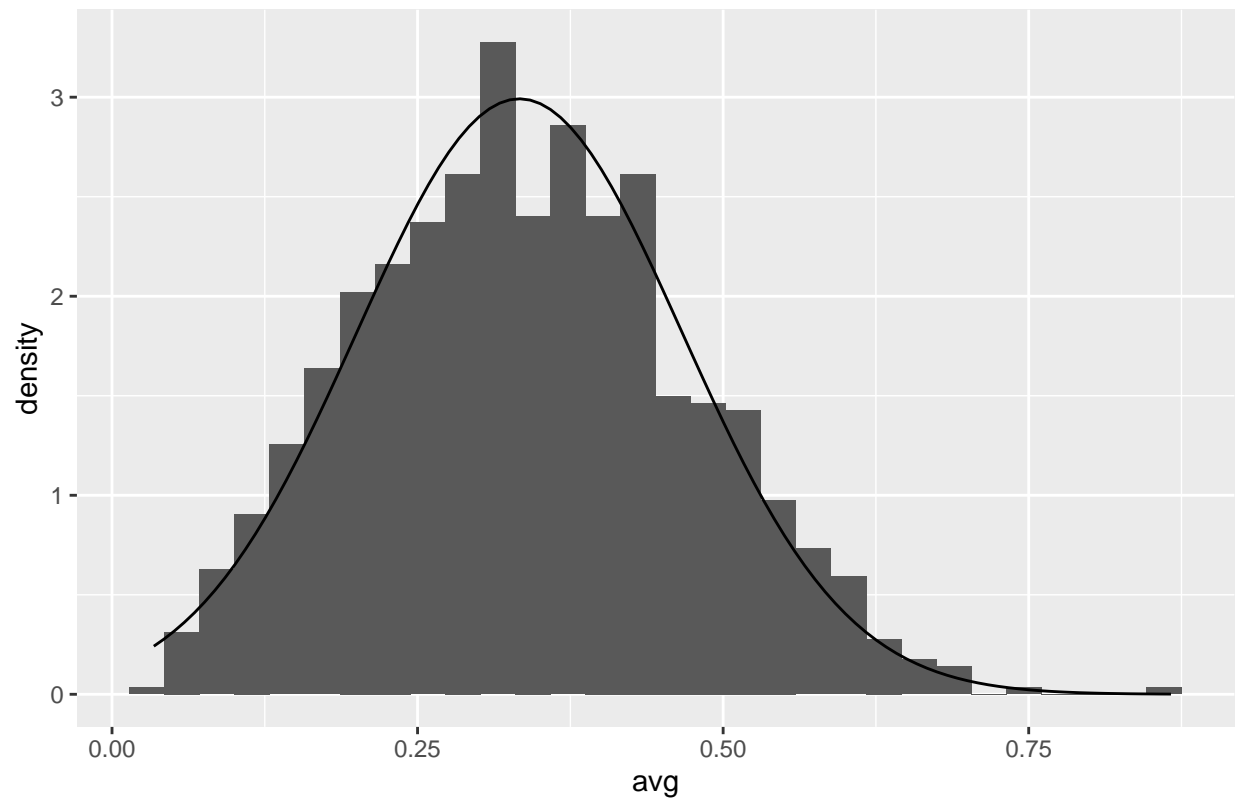
```
library(ggplot2)
f <- function(alpha,n){                                     # a function to calculate means
  m <- matrix(0,ncol = n,nrow = 1000)
  for (i in 1:1000) {
    m[i,] <- rbeta(n, shape1 = alpha, shape2 = 1, ncp = 0)
  }
  return(rowMeans(m))
}

alpha <- 0.5                                                #parameters (0.5,5)
n <- 5
avg <- f(alpha,n)

ggplot(data = as.data.frame(avg),aes(x=avg)) + geom_histogram(aes(y = ..density..)) +
  stat_function(fun = dnorm, args = list(mean = alpha/(alpha+1), sd = sqrt(alpha*1/((alpha + 1)^2 * (alpha + 1))))

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

alpha=0.5, n=5

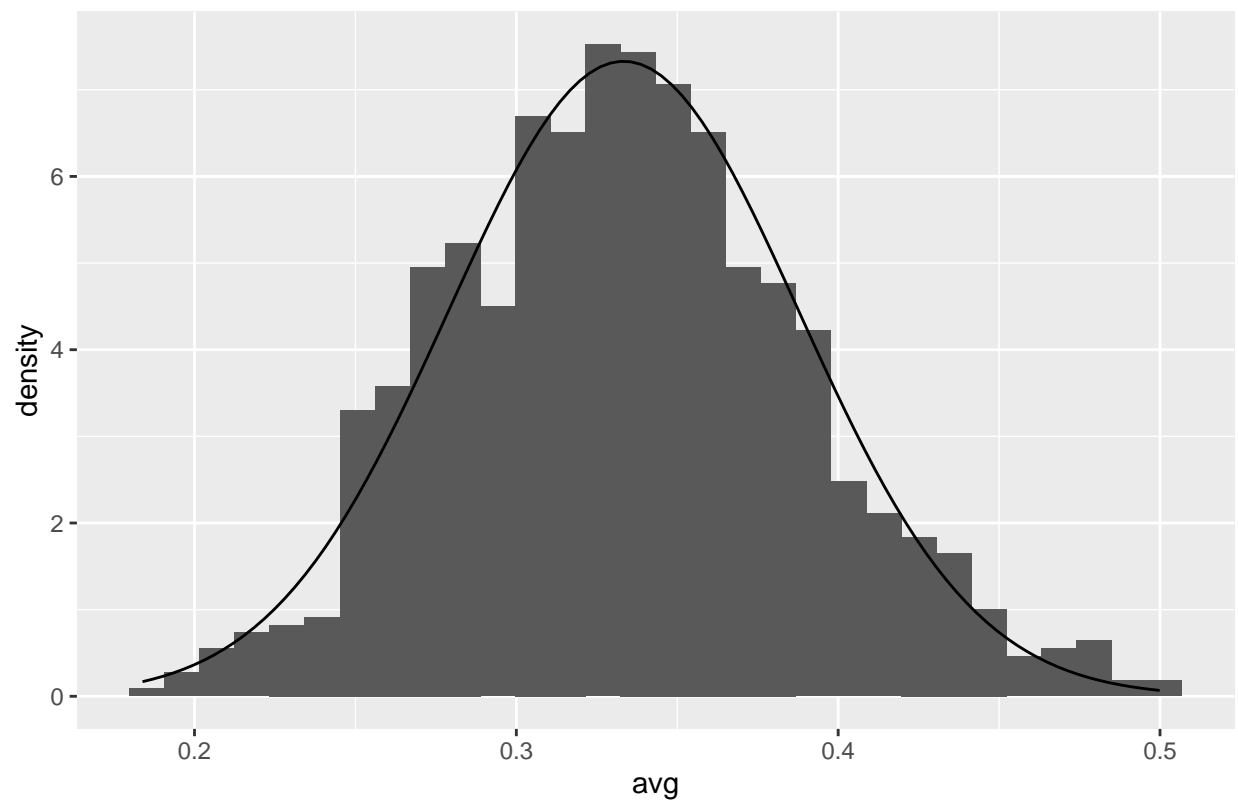


```
alpha <- 0.5                                     #parameters (0.5,30)
n <- 30
avg <- f(alpha,n)

ggplot(data = as.data.frame(avg),aes(x=avg)) + geom_histogram(aes(y = ..density..)) +
  stat_function(fun = dnorm, args = list(mean = alpha/(alpha+1), sd = sqrt(alpha*1/((alpha + 1)^2 * (alpha + 1))))

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

alpha=0.5, n=30

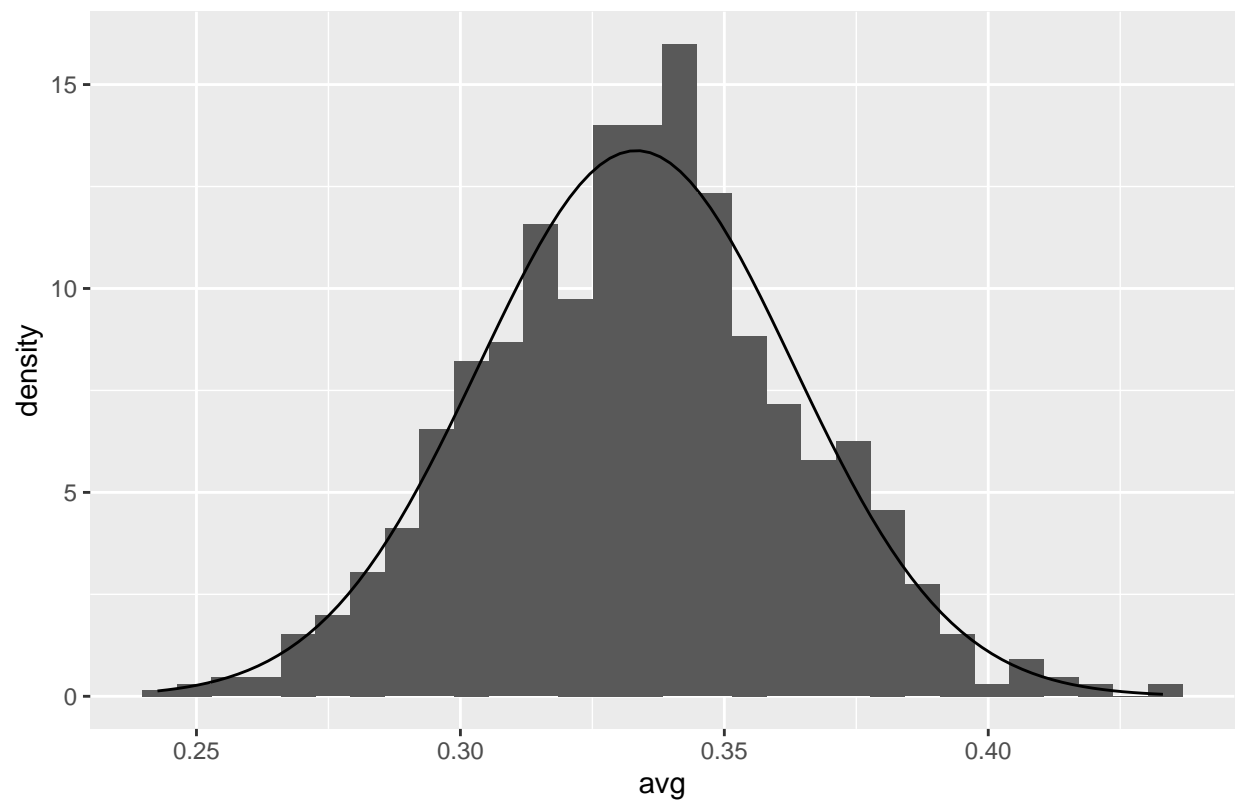


```
alpha <- 0.5 #parameters (0.5,100)
n <- 100
avg <- f(alpha,n)

ggplot(data = as.data.frame(avg),aes(x=avg)) + geom_histogram(aes(y = ..density..)) +
  stat_function(fun = dnorm, args = list(mean = alpha/(alpha+1), sd = sqrt(alpha*1/((alpha + 1)^2 * (alpha + 1))))

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

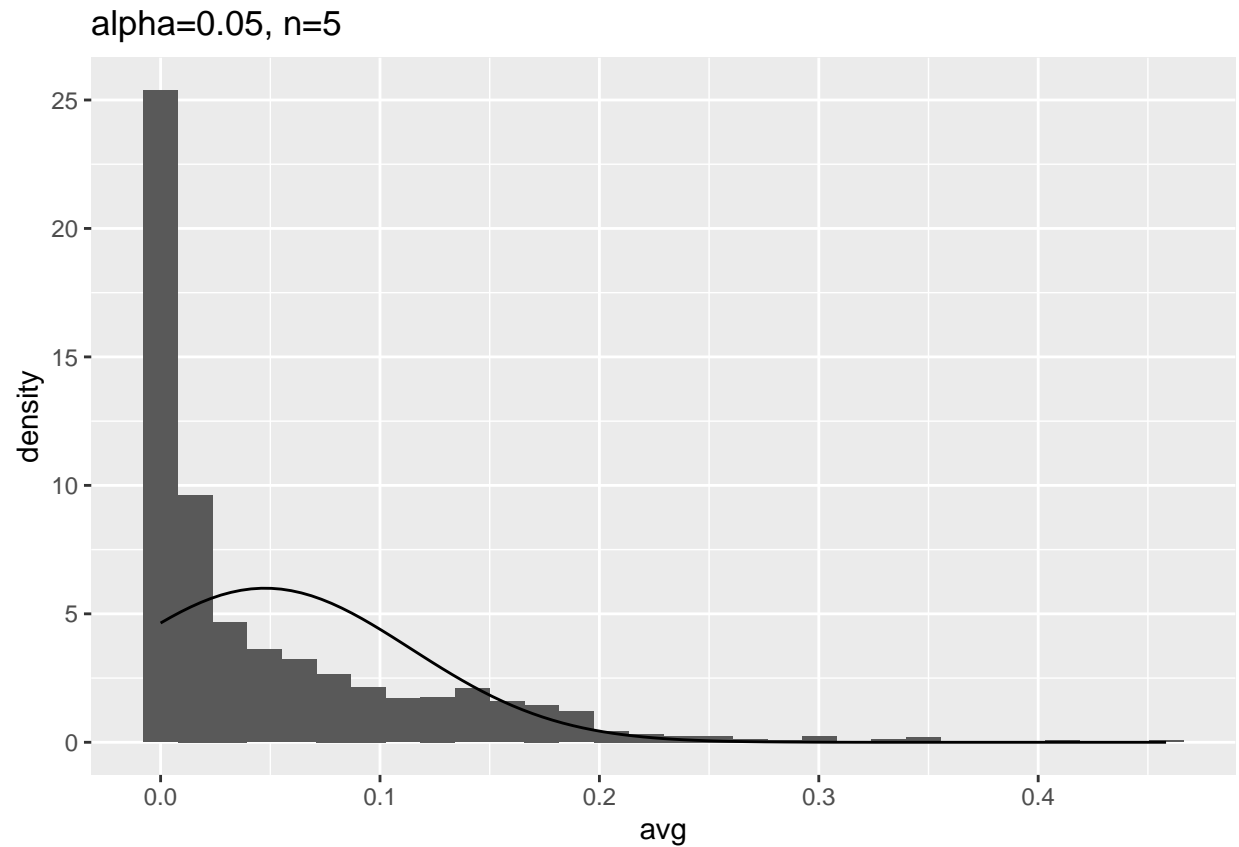
alpha=.5, n=100



```
alpha <- 0.05                                     #parameters (0.05,5)
n <- 5
avg <- f(alpha,n)

ggplot(data = as.data.frame(avg), aes(x=avg)) + geom_histogram(aes(y = ..density..)) +
  stat_function(fun = dnorm, args = list(mean = alpha/(alpha+1), sd = sqrt(alpha*1/((alpha + 1)^2 * (alpha + 1))))

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

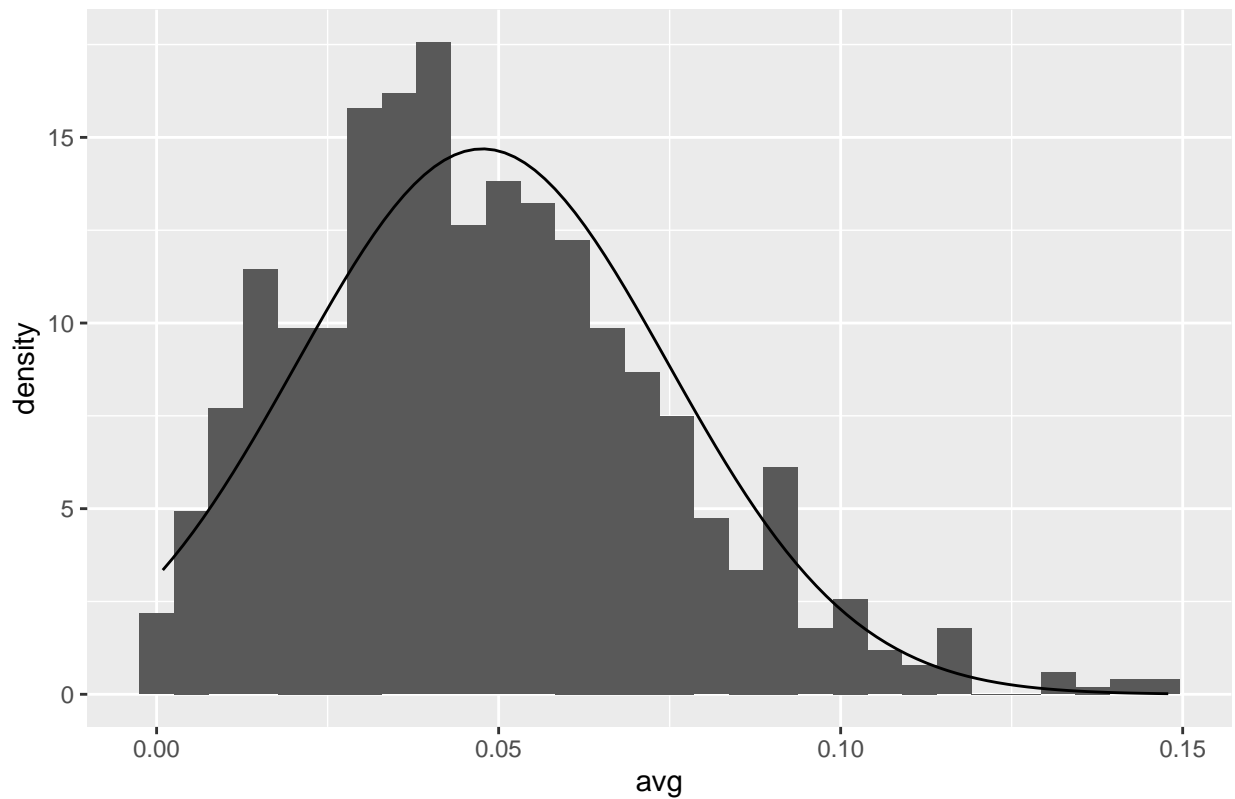


```
alpha <- 0.05                                     #parameters (0.05,30)
n <- 30
avg <- f(alpha,n)

ggplot(data = as.data.frame(avg),aes(x=avg)) + geom_histogram(aes(y = ..density..)) +
  stat_function(fun = dnorm, args = list(mean = alpha/(alpha+1), sd = sqrt(alpha*1/((alpha + 1)^2 * (alpha + 1))))

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

alpha=0.05, n=30

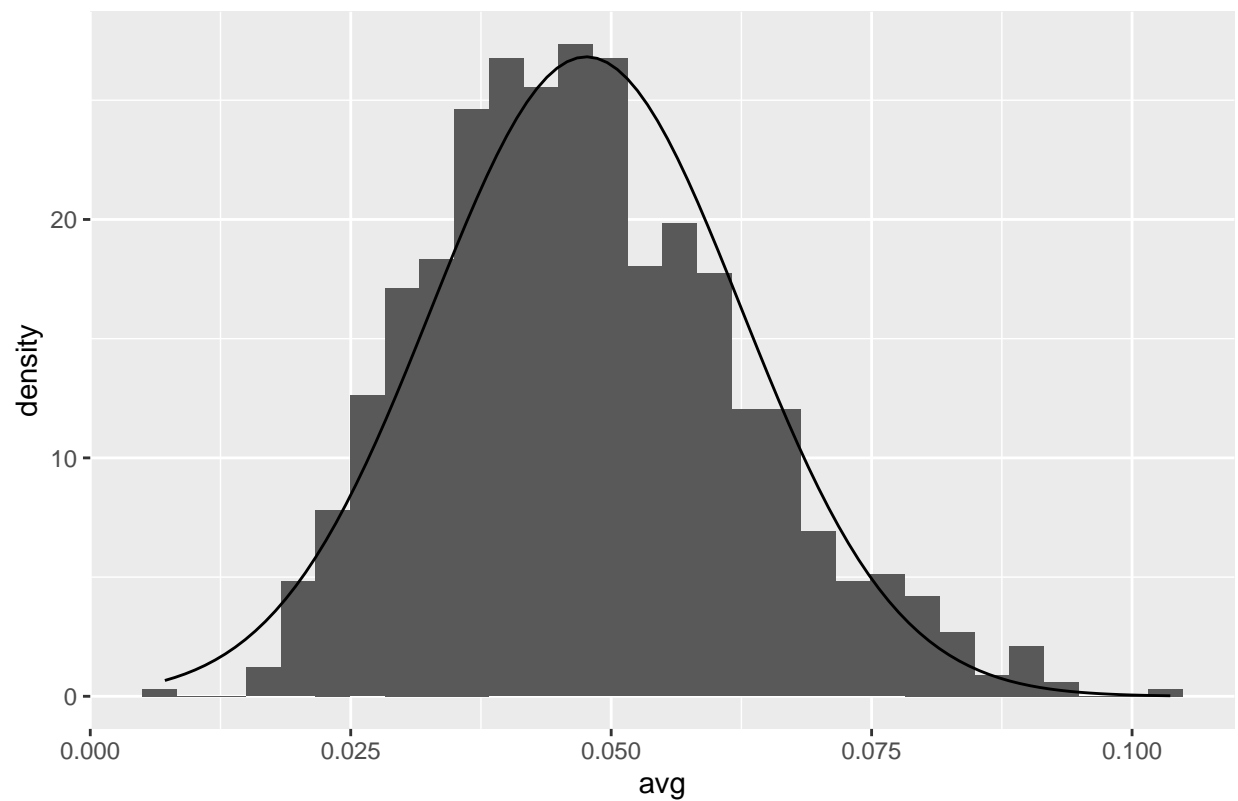


```
alpha <- 0.05                                     #parameters (.05,30)
n <- 100
avg <- f(alpha,n)

ggplot(data = as.data.frame(avg), aes(x=avg)) + geom_histogram(aes(y = ..density..)) +
  stat_function(fun = dnorm, args = list(mean = alpha/(alpha+1), sd = sqrt(alpha*1/((alpha + 1)^2 * (alpha + 1))))

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

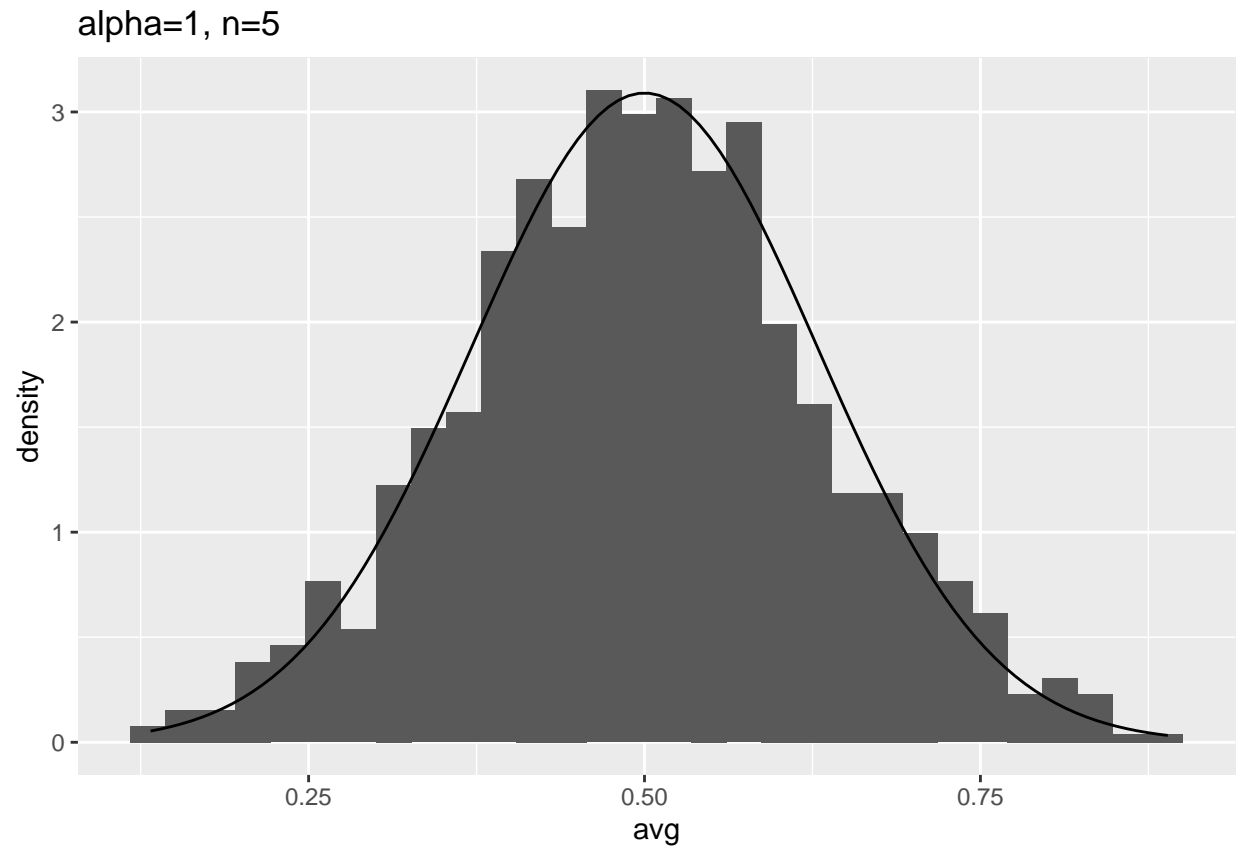
alpha=.05, n=100



```
alpha <- 1 #parameters (1,5)
n <- 5
avg <- f(alpha,n)

ggplot(data = as.data.frame(avg),aes(x=avg)) + geom_histogram(aes(y = ..density..)) +
  stat_function(fun = dnorm, args = list(mean = alpha/(alpha+1), sd = sqrt(alpha*1/((alpha + 1)^2 * (alpha + 1))))

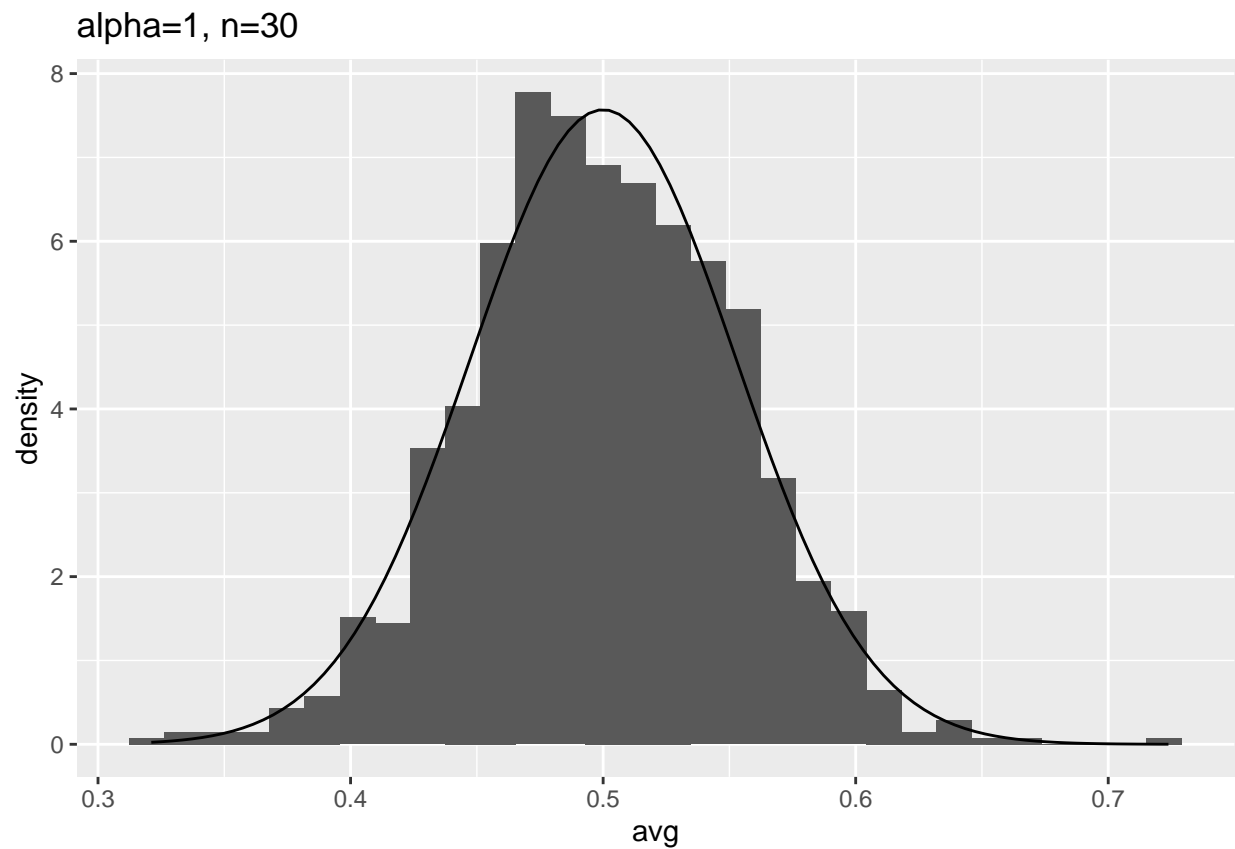
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
alpha <- 1                                     #parameters (1,30)
n <- 30
avg <- f(alpha,n)

ggplot(data = as.data.frame(avg),aes(x=avg)) + geom_histogram(aes(y = ..density..)) +
  stat_function(fun = dnorm, args = list(mean = alpha/(alpha+1), sd = sqrt(alpha*1/((alpha + 1)^2 * (alpha + 1))))

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

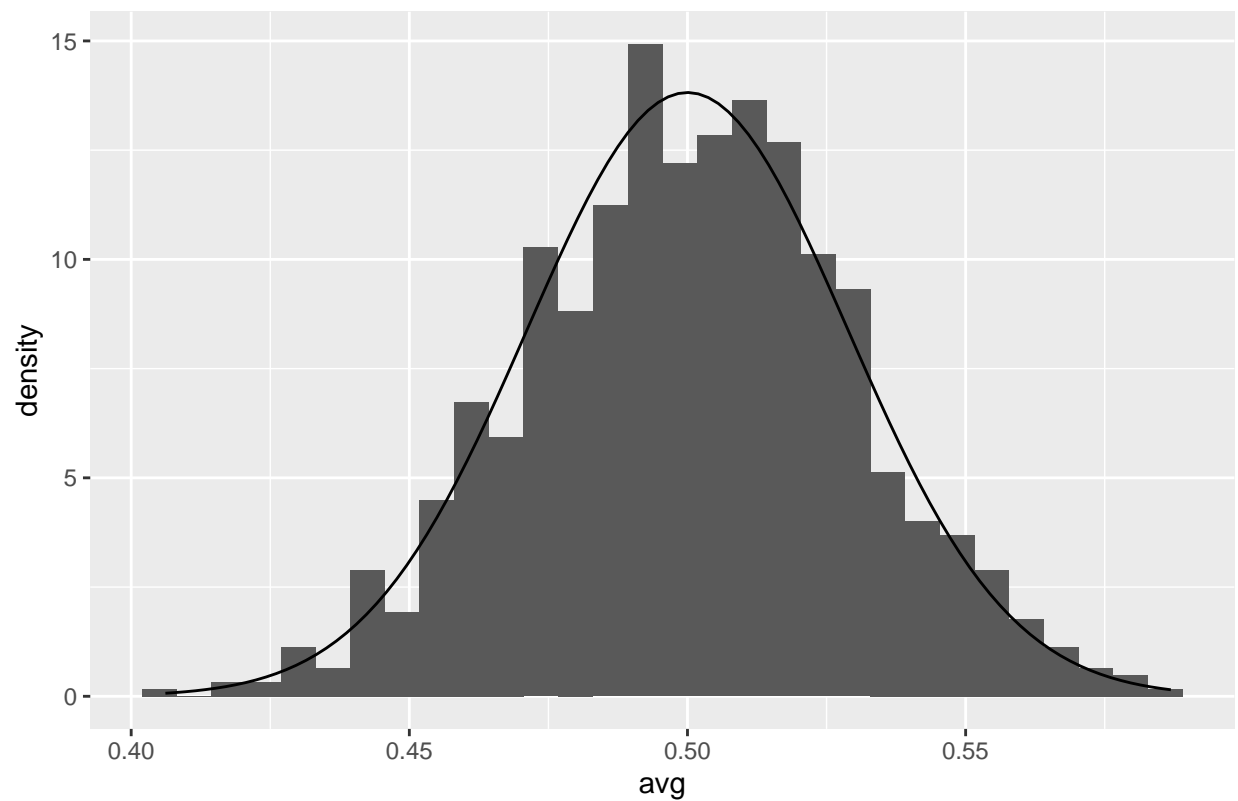



```
alpha <- 1 #parameters (1,100)
n <- 100
avg <- f(alpha,n)

ggplot(data = as.data.frame(avg),aes(x=avg)) + geom_histogram(aes(y = ..density..)) +
  stat_function(fun = dnorm, args = list(mean = alpha/(alpha+1), sd = sqrt(alpha*1/((alpha + 1)^2 * (alpha + 1))))

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

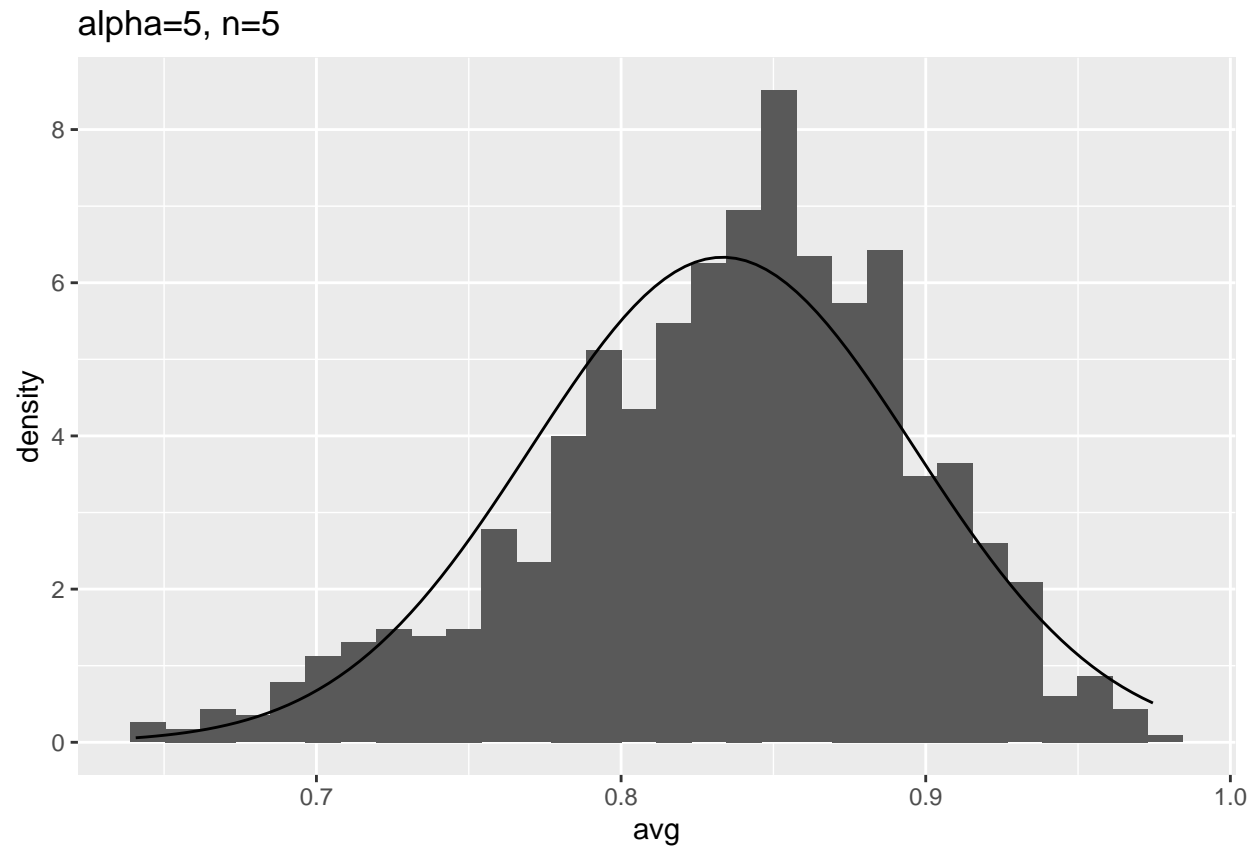
alpha=1, n=100



```
alpha <- 5                                     #parameters (5,5)
n <- 5
avg <- f(alpha,n)

ggplot(data = as.data.frame(avg),aes(x=avg)) + geom_histogram(aes(y = ..density..)) +
  stat_function(fun = dnorm, args = list(mean = alpha/(alpha+1), sd = sqrt(alpha*1/((alpha + 1)^2 * (alpha + 1))))

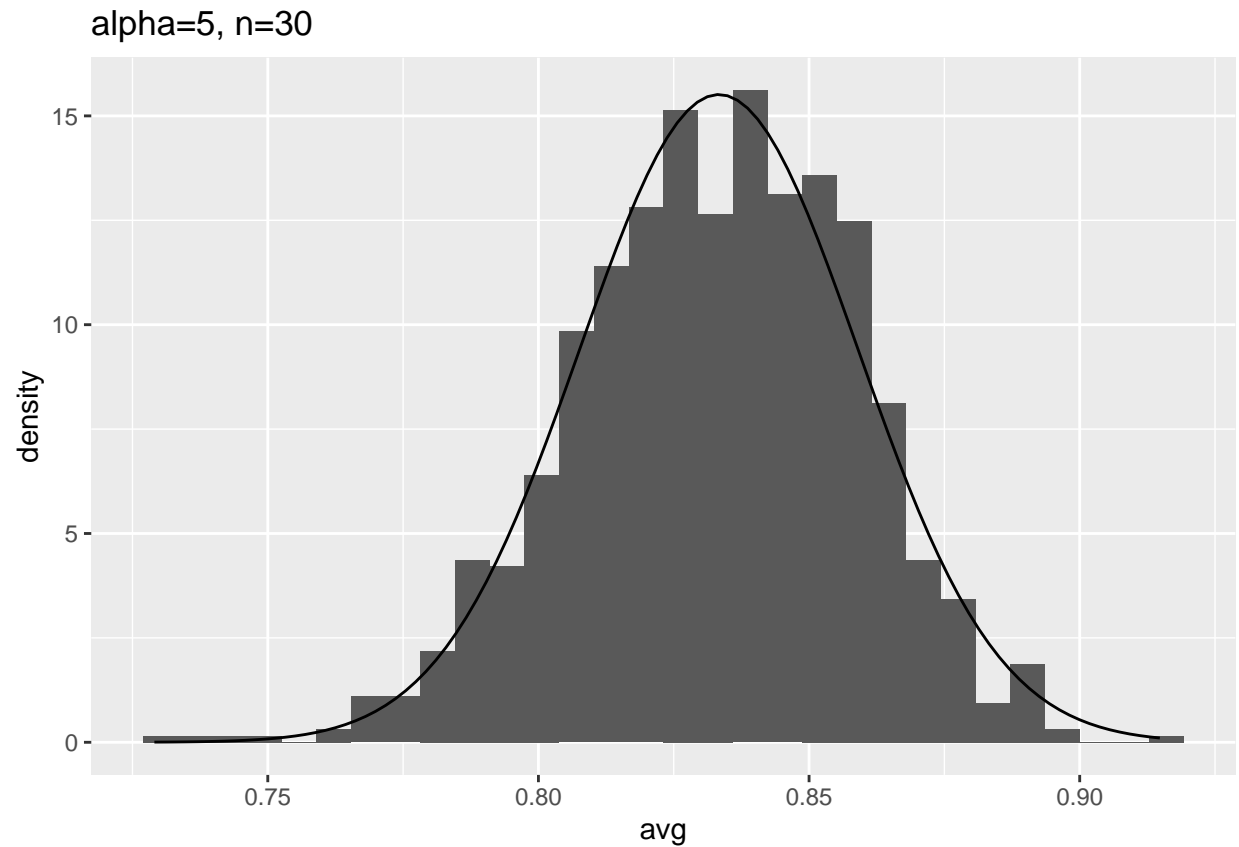
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
alpha <- 5                                     #parameters (5,30)
n <- 30
avg <- f(alpha,n)

ggplot(data = as.data.frame(avg),aes(x=avg)) + geom_histogram(aes(y = ..density..)) +
  stat_function(fun = dnorm, args = list(mean = alpha/(alpha+1), sd = sqrt(alpha*1/((alpha + 1)^2 * (alpha + 1))))

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
alpha <- 5                                     #parameters (5,100)
n <- 100
avg <- f(alpha,n)

ggplot(data = as.data.frame(avg),aes(x=avg)) + geom_histogram(aes(y = ..density..)) +
  stat_function(fun = dnorm, args = list(mean = alpha/(alpha+1), sd = sqrt(alpha*1/((alpha + 1)^2 * (alpha + 1))))

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

