

# ST501 Group Project

*Our Group*

*7/20/2018*

## Introduction

## Part 1 Convergences

### A

$$f_Y(y) = \frac{1}{2b} e^{-\left(\frac{|y-\mu|}{b}\right)}$$

Given:  $\mu = 0$  and  $b = 5$

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} y * f(y) dy \\ &= \int_{-\infty}^0 y * \frac{1}{10} e^{\frac{y}{5}} + \int_0^{\infty} y * \frac{1}{10} e^{\frac{-y}{5}} \\ &= \frac{1}{10} \left[ \int_{-\infty}^0 y e^{\frac{y}{5}} + \int_0^{\infty} y * e^{\frac{-y}{5}} \right] \\ &= \frac{1}{10} \left[ \left[ (5y - 25) e^{\frac{y}{5}} \right]_{-\infty}^0 + \left[ (-5y - 25) e^{\frac{-y}{5}} \right]_0^{\infty} \right] \\ &= \frac{1}{10} [-25 + 25] \\ E(Y) &= 0 \blacksquare \\ E(Y^2) &= \int_{-\infty}^{\infty} y^2 * f(y) dy \\ &= \frac{1}{10} \left[ \int_{-\infty}^0 y^2 e^{\frac{y}{5}} + \int_0^{\infty} y^2 * e^{\frac{-y}{5}} \right] \\ &= \frac{1}{10} * [250 + 250] \\ E(Y^2) &= 50 \blacksquare \end{aligned}$$

Therefor  $E(Y^2)$  exists.

Thus:

$E(Y) = 0$ ,  $E(Y^2) = 50$  then  $Var(Y) = 50$  by the variance computing formula.

$$L = \frac{1}{n} \sum_{i=1}^n Y_i^2$$

By the Law of Large Numbers:

$$\frac{1}{n} \sum_{i=1}^n Y_i \xrightarrow{p} E(Y_i)$$

By generalisation of the Law of Large Numbers:

$$L = \frac{1}{n} \sum_{i=1}^n Y_i^2 \xrightarrow{p} E(Y_i^2) = 50$$

Therefore:

$$L \xrightarrow{p} 50 \quad \blacksquare$$

## B

By the continuity theorem,  $K = \sqrt{L} \xrightarrow{p} \sqrt{50}$

Thus:  $K \xrightarrow{p} \sqrt{50}$ .

## C

Derive the CDF of the double exponential distribution.

$$f_Y(y) = \frac{1}{2b} * e^{-(\frac{|y-\mu|}{b})}$$

This yields two cases for the CDF:

$$F_Y(y) = \begin{cases} \int_{-\infty}^t f_Y(Y) dt & \text{for } y \leq \mu \\ \int_t^{\infty} f_Y(Y) dt & \text{for } y \geq \mu \end{cases}$$

Solving for these leaves us with the following:

For

$$y \leq \mu$$

$$\int_{-\infty}^y f_Y(Y) dt$$

$$\int_{-\infty}^y \frac{1}{2b} * e^{-(\frac{t-\mu}{b})} dt$$

We know that in this case we are solving for when the exponent is positive.

$$\int_{-\infty}^y \frac{1}{2b} * e^{-(\frac{t-\mu}{b})} dt$$

$$\frac{1}{2b} * e^{(\frac{t-\mu}{b})} * b|_{-\infty}^y = \frac{1}{2} e^{(\frac{t-\mu}{b})} - 0 = \frac{1}{2} e^{(\frac{t-\mu}{b})}$$

For the case where  $y \geq \mu$

$$\int_{-\infty}^y f_Y(Y) dt$$

$$\int_{-\infty}^y \frac{1}{2b} * e^{-(\frac{t-\mu}{b})} dt$$

This must be further split into:

$$\int_{-\infty}^{\mu} \frac{1}{2b} * e^{-(\frac{t-\mu}{b})} dt + \int_{\mu}^y \frac{1}{2b} * e^{-(\frac{t-\mu}{b})} dt$$

$$\frac{1}{2b} * e^{(-\frac{t+\mu}{b})} * -b|_{-\infty}^{\mu} + \frac{1}{2b} * e^{(-\frac{t+\mu}{b})} * -b|_{\mu}^y$$

$$\frac{1}{2} - \frac{1}{2} e^{\frac{\mu-y}{b}} + \frac{1}{2}$$

Thus for this case:

$$1 - \frac{1}{2} e^{\frac{\mu-y}{b}}$$

In conclusion the CDF for the double exponential distribution is:

$$F_Y(y) = \begin{cases} \frac{1}{2} e^{(\frac{y-\mu}{b})} & \text{for } y \leq \mu \\ 1 - \frac{1}{2} e^{\frac{\mu-y}{b}} & \text{for } y \geq \mu \end{cases}$$

## D and E

First to generate these random values we need to create the inverse functions.

For the case when:

$$y < \mu$$

$$u = \frac{1}{2} e^{\frac{y-\mu}{b}}$$

$$b * \log(2u) = y - \mu$$

Thus

$$y = \mu + b * \log(2u)$$

for

$$u > 0$$

$$y < \mu$$

For the case when

$$y \geq \mu$$

$$u = 1 - \frac{1}{2}e^{-\frac{y-\mu}{b}}$$
$$-b * \log(2 - 2u) = y - \mu$$

Thus

$$\mu - b * \log(2 - 2u) = y$$

For  $u < 1$ ,  $x \geq \mu$

Using these inverse functions we can now do simulations.

```
set.seed(336)

test_func<-function(u){
  out_come_1 <- 0 + 5*log(2*u)
  out_come_2 <- 0 - 5*log(2-2*u)
  cbind(out_come_1^2, out_come_2^2)
}

a<-(test_func(seq(0.01,.999,.001)))

rdoublex <- function(u, mu = 0, b = 5){
  if( u < 0.5){
    out_come_1 <- mu + b*log(2*u)
    out_come_1
  } else{
    out_come_2 <- mu - b*log(2-2*u)
    out_come_2
  }
}

output_matrix <- matrix(0, nrow= 250, ncol = 50)

# Basic Function
for( j in 1:50){
  for(i in 1:250){
    random_value <- runif(i, 0, 1)
    output_matrix[i, j]<- mean(vapply(random_value, FUN = rdoublex, double(1)))
  }
}

p1 <- output_matrix %>%
  as.data.frame() %>%
  mutate(N = 1:250) %>%
  gather(replication, value, -N) %>%
  ggplot(aes(N, value))+
  geom_point(alpha=1/5)+
  theme_minimal()+
```

```

labs(
  title = "Convergence of a Laplace Distribution with  $\mu = 0$ ,  $b = 5$ ",
  subtitle = "50 Samples Drawn per Sample Size, N",
  x = "Sample Size, N"
)

#L

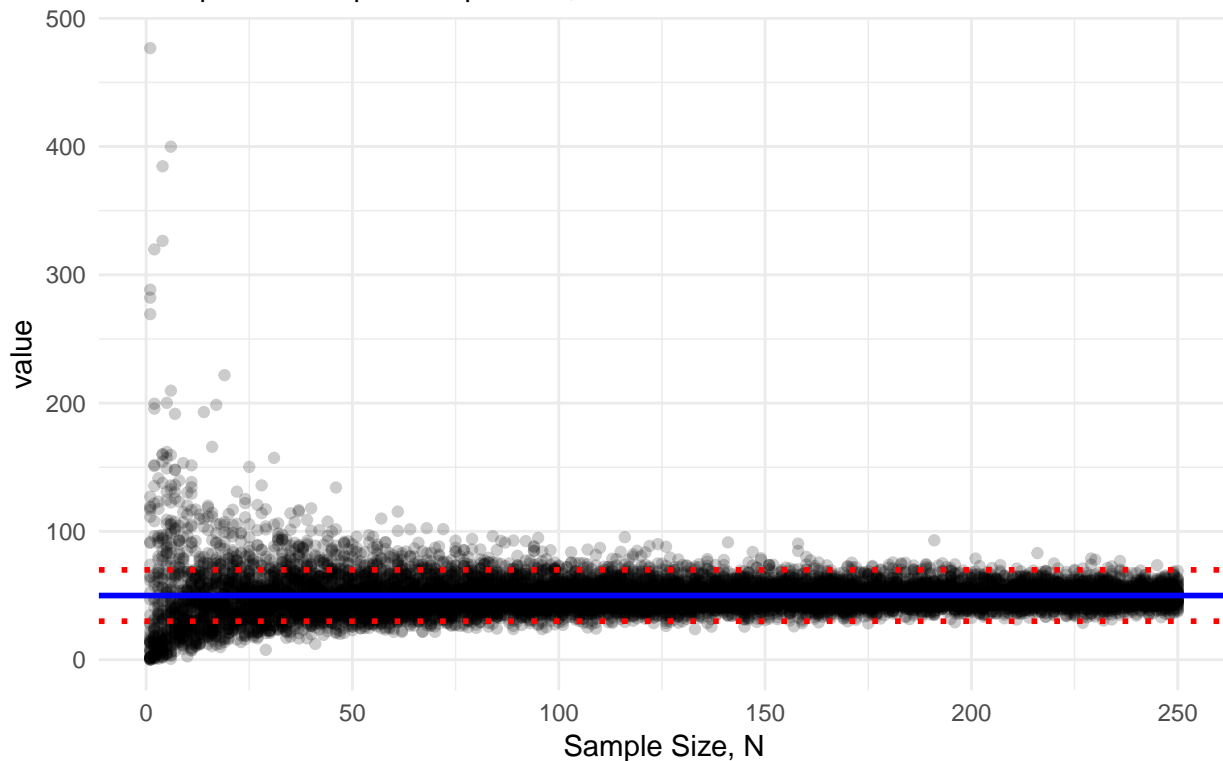
output_matrix <- matrix(0, nrow= 250, ncol = 50)

# Basic Function
for( j in 1:50){
  for(i in 1:250){
    random_value <- runif(i, 0, 1)
    output_matrix[i, j]<- mean(vapply(random_value, FUN = rdouplex, double(1))^2)
  }
}

limit_one <- 20
p2 <- output_matrix %>%
  as.data.frame() %>%
  mutate(N = 1:250) %>%
  gather(replication, value, -N) %>%
  ggplot(aes(N, value))+
  geom_point(alpha=1/5)+
  theme_minimal()+
  labs(
    title = "Convergence of a Laplace Distribution of  $Y_i^2$  with  $\mu = 0$ ,  $b = 5$ ",
    subtitle = "50 Samples Drawn per Sample Size, N",
    x = "Sample Size, N"
  )+
  geom_hline(yintercept = 50, size = 1, color = "blue")+
  geom_hline(yintercept = 50 + limit_one, linetype = "dotted",
    color = "red", size = 1)+
  geom_hline(yintercept = 50 - limit_one,
    linetype = "dotted", color = "red", size = 1)
p2

```

Convergence of a Laplace Distribution of  $Y_i^2$  with  $\mu = 0$ ,  $b = 5$   
 50 Samples Drawn per Sample Size, N



```
#K

output_matrix <- matrix(0, nrow= 250, ncol = 50)

# Basic Function
for( j in 1:50){
  for(i in 1:250){
    random_value <- runif(i, 0, 1)
    output_matrix[i, j]<- sqrt(mean(vapply(random_value, FUN = rdouplex, double(1))^2))
  }
}

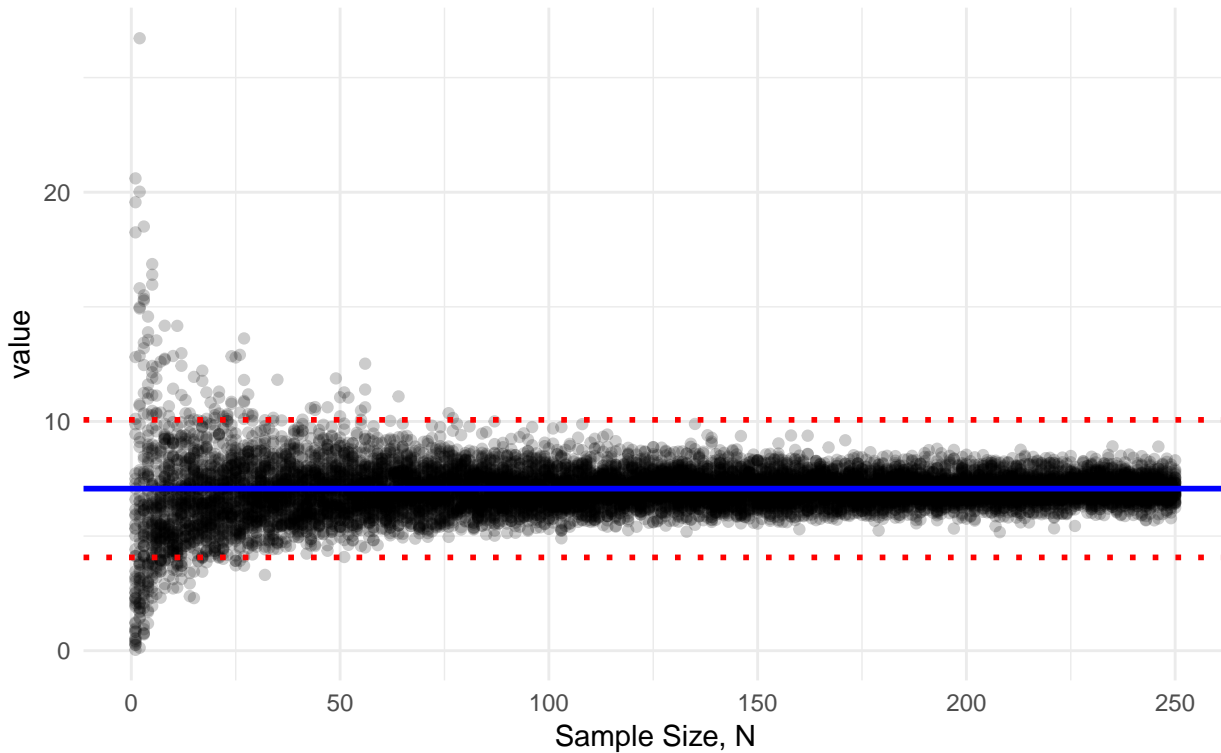
limit_two <- 3
p3 <- output_matrix %>%
  as.data.frame() %>%
  mutate(N = 1:250) %>%
  gather(replication, value, -N) %>%
  ggplot(aes(N, value))+
  geom_point(alpha=1/5)+
  theme_minimal()+
  labs(
    title = "Convergence of a Laplace Distribution of sqrt(Y_i^2) with mu= 0, b = 5",
    subtitle = "50 Samples Drawn per Sample Size, N",
    x = "Sample Size, N"
  )+
  geom_hline(yintercept = 7.07, size = 1, color = "blue")+
  geom_hline(yintercept = 7.07 + limit_two,
```

```

linetype = "dotted", color = "red", size = 1)+
geom_hline(yintercept = 7.07 - limit_two,
linetype = "dotted", color = "red", size = 1)
p3

```

Convergence of a Laplace Distribution of  $\sqrt{Y_i^2}$  with  $\mu=0$ ,  $b=5$   
 50 Samples Drawn per Sample Size, N



These plots show that as the sample size increases the values we see approach the expected values. This is graphically showing the convergence of these functions to their expected values as the sample size increases.

## Part 2

```

#Simulate 10000 samples of size 10,50,100,250 from Laplace distribution
n<-c(10,50,100,250)
N<-50
mu<-0
beta<-5

rlaplace <- function(n = 1, mu = 0, beta = 5) {
  q <- runif(n)
  ifelse(q < 0.5, beta * log(2 * q) + mu, -beta * log(2 * (1 - q)) + mu)
}

#list to save data values in
data<-list()
for(i in 1:length(n)){data[[i]]<-matrix(0,nrow=N,ncol=n[i])}
#Create the data

```

```

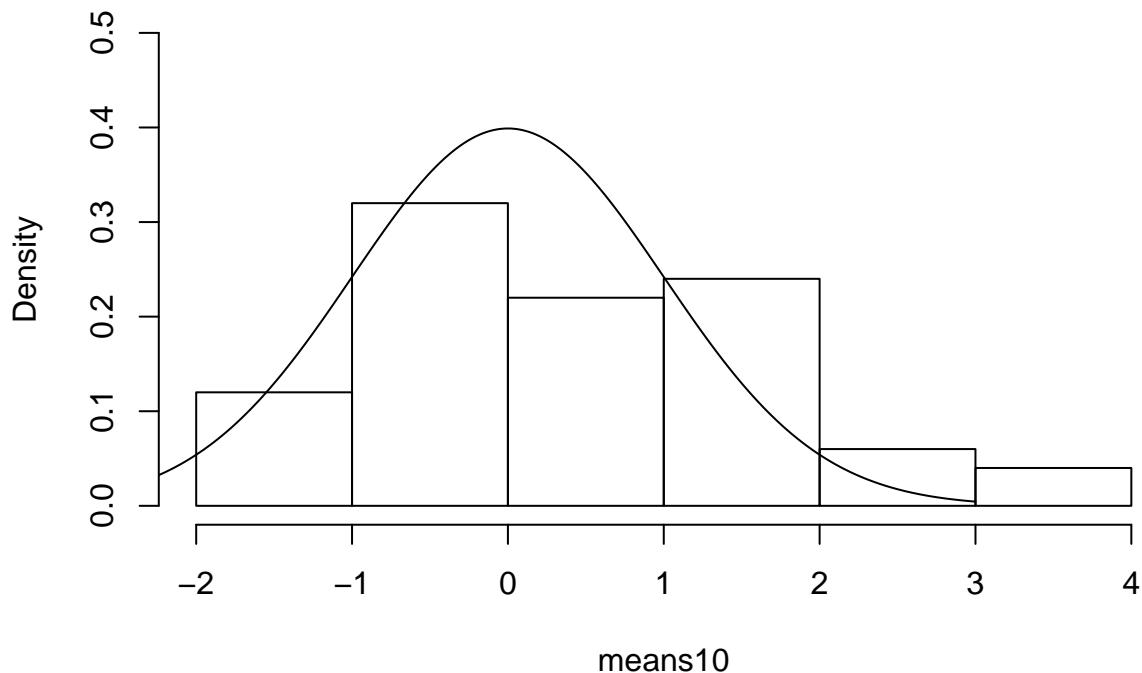
#loop over sample sizes
for (j in 1:length(n)){
  #loop over data sets
  for (i in 1:N){
    data[[j]][i,]<-rlaplace(n=n[j],0,5)
  }
}

#calculate the z statistic for each sample
means10<-apply(X=data[[1]],FUN=function(data){(mean(data)-mu)/(beta/sqrt(n[1]))},MARGIN=1)
means50<-apply(X=data[[2]],FUN=function(data){(mean(data)-mu)/(beta/sqrt(n[2]))},MARGIN=1)
means100<-apply(X=data[[3]],FUN=function(data){(mean(data)-mu)/(beta/sqrt(n[3]))},MARGIN=1)
means250<-apply(X=data[[3]],FUN=function(data){(mean(data)-mu)/(beta/sqrt(n[4]))},MARGIN=1)

# Make Histogram
hist(means10,main=paste("Histogram of laplace's with n=",n[1]," from
                        N(",mu,"",beta^2,")",sep=""), ylim=c(0,.5), prob=T)
lines(seq(from=-3,to=3,by=0.01),dnorm(seq(from=-3,to=3,by=0.01)))

```

**Histogram of laplace's with n=10 from  
N(0,25)**



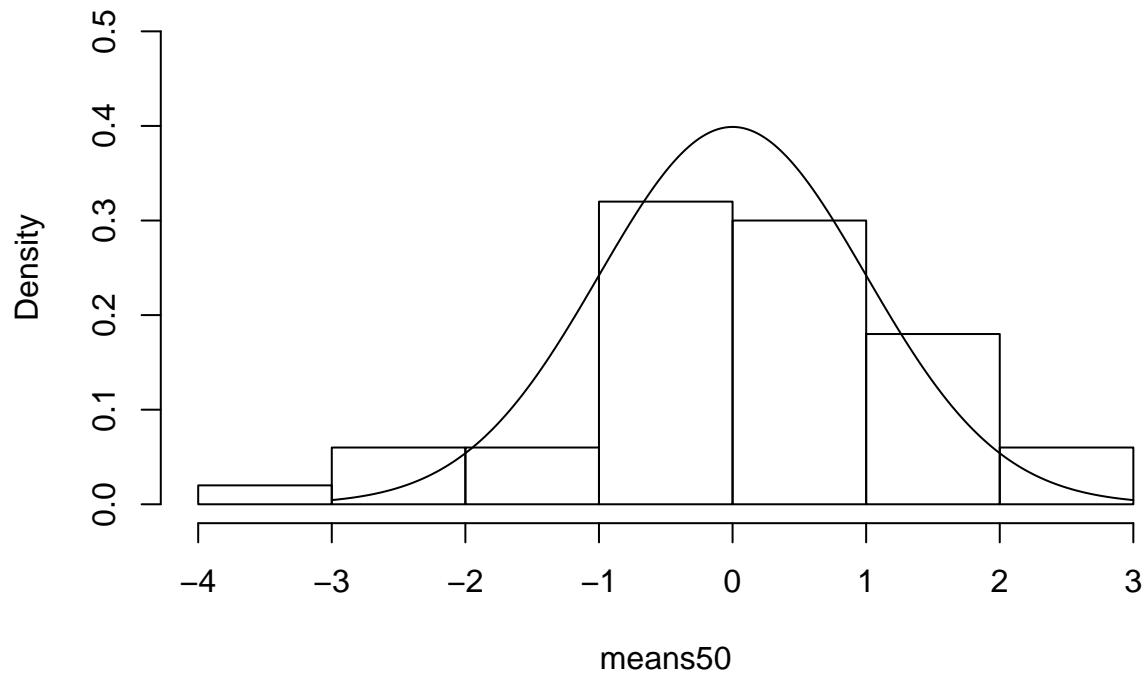
```

hist(means50,main=paste("Histogram of laplace's with n=",n[2]," from
                        N(",mu,"",beta^2,")",sep=""),prob=T,ylim=c(0,.5),)
lines(seq(from=-3,to=3,by=0.01),dnorm(seq(from=-3,to=3,by=0.01)))

```

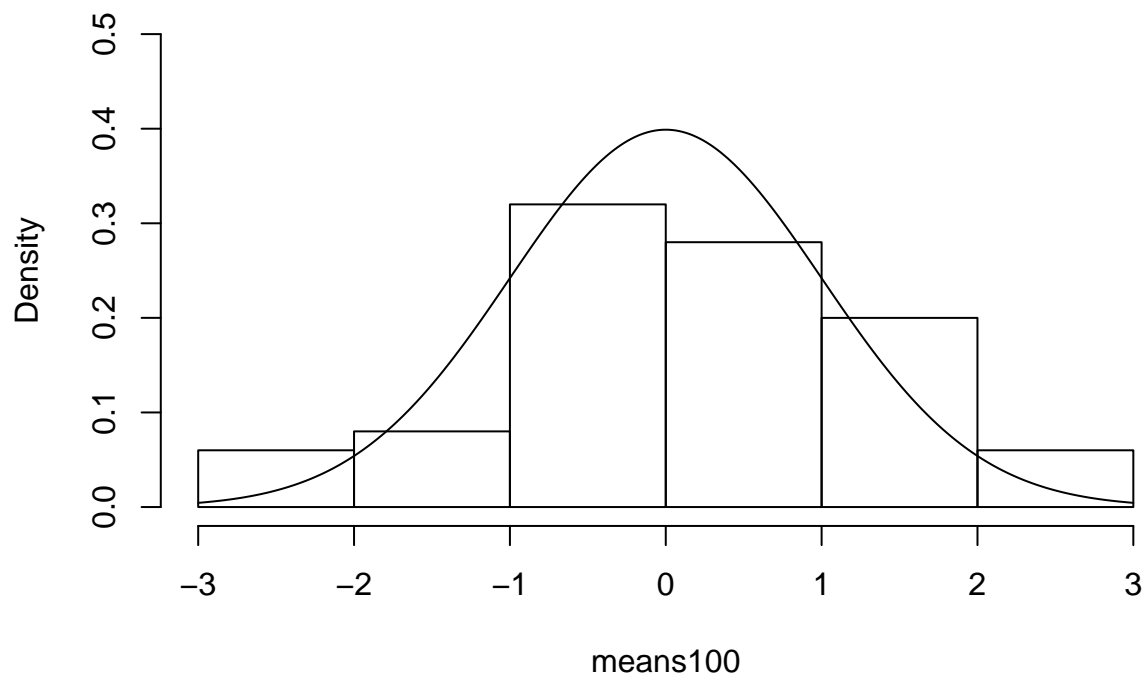


**Histogram of laplace's with n=50 from  
N(0,25)**



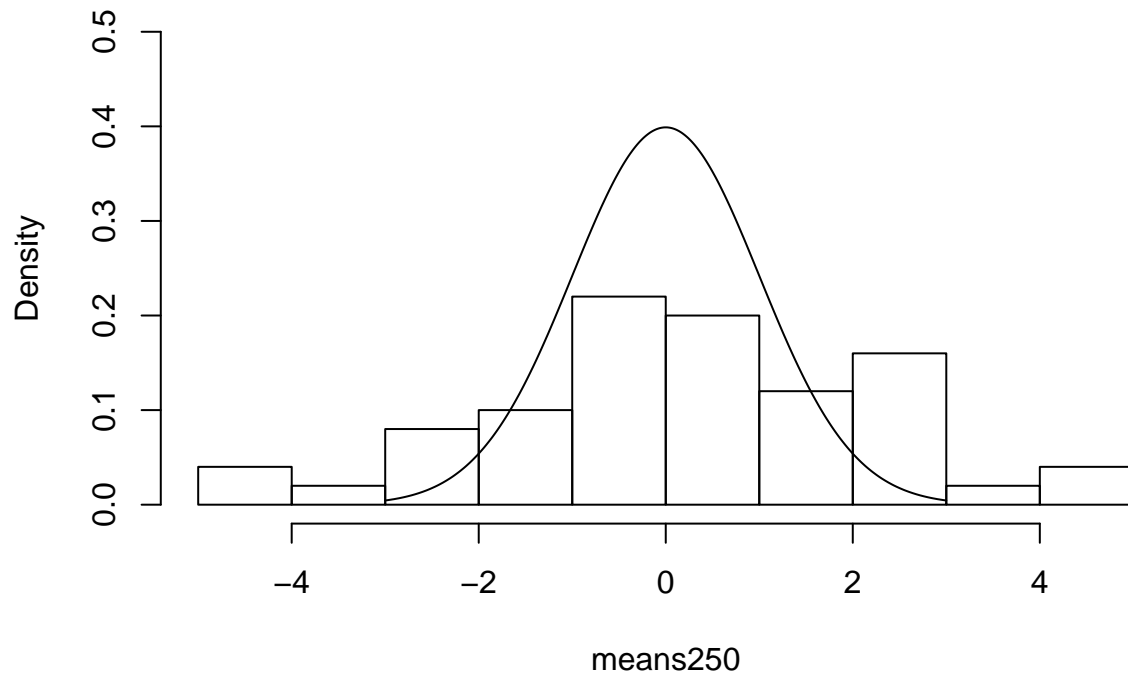
```
hist(means100,main=paste("Histogram of laplace's with n=",n[3]," from
                          N(",mu," ",beta^2,")",sep=""),prob=T,ylim=c(0,.5),)
lines(seq(from=-3,to=3,by=0.01),dnorm(seq(from=-3,to=3,by=0.01)))
```

**Histogram of laplace's with n=100 from  
N(0,25)**



```
hist(means250,main=paste("Histogram of laplace's with n=",n[4]," from
                          N(",mu,",",beta^2,")",sep=""),prob=T,ylim=c(0,.5),)
lines(seq(from=-3,to=3,by=0.01),dnorm(seq(from=-3,to=3,by=0.01)))
```

**Histogram of laplace's with n=250 from  
N(0,25)**



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

- A
- B
- C
- D