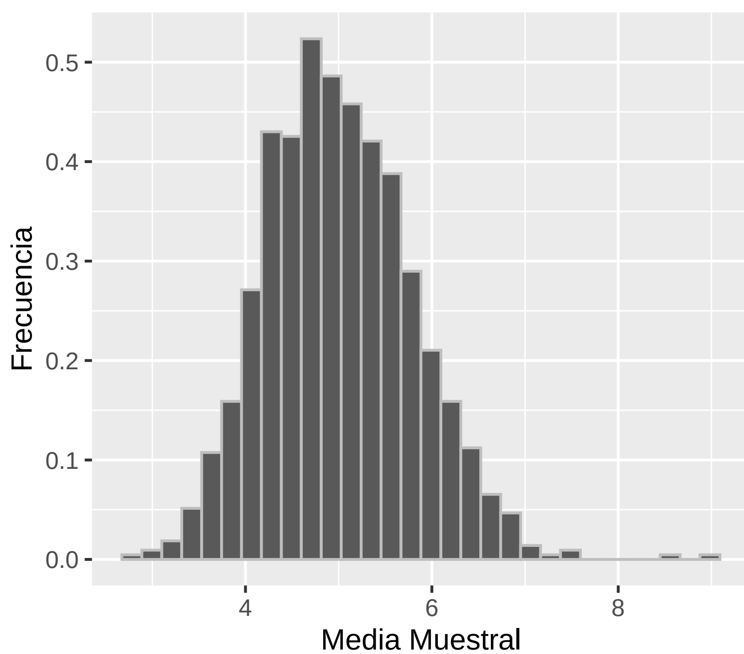
Statistical Inference Course Project

luis manuel

Sumulation of the central limit theorem

We are going to verify the Central Limit Terema which states that when the sample size is large enough, the sample distribution of the average follows a Normal with mean μ and variance $\frac{\sigma^2}{n}$, to verify this we will obtain 1000 simulations of 40 exponential variables with $\lambda = 0.2$

```
library(ggplot2)
set.seed(650)
mns<-NULL
for (i in 1 : 1000) mns = c(mns, mean(rexp(40,0.2)))
ggplot(data.frame(M=mns), aes(x=M)) + geom_histogram(aes(y=..density..),color="grey")+xlab("Media Muest
```



We can see that it has a shape similar to that of a normal distribution and according to the theory it is distributed as:

$$\bar{X} \sim N(\mu_{\bar{x}} = \mu, \sigma^2_{\bar{x}} = \frac{\sigma^2}{n})$$

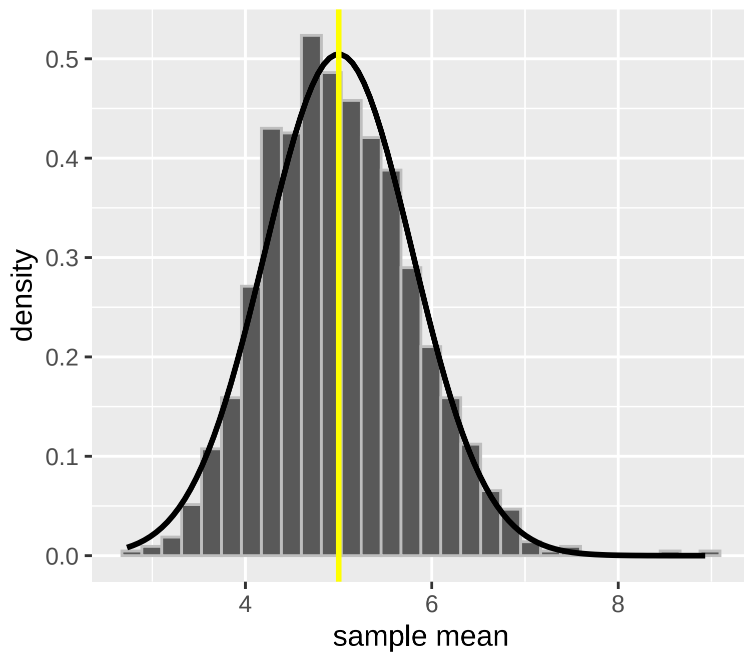
where $\mu = \frac{1}{\lambda} = 5$ and $\sigma^2 = \frac{1}{\lambda}^2 = 25$, then the parameters of the distribution of the mean should be $\mu_{\bar{x}} = \mu = \frac{1}{\lambda} = 5$ and $\sigma^2_{\bar{x}} = \frac{\sigma^2}{n} = 0.625$ to verify this let us calculate the mean and variance of the sampling distribution of the mean

$$\mu_{\bar{x}}=5.0097214$$

$$\sigma^2_{\bar{x}}=0.617403$$

we can realize that the values are not exactly the same but they are very approximate, now let's compare the distribution, the black curve is a normal distribution with mean $\mu = 5$ and variance $\sigma = 0.625$ and the yellow line corresponds to 5

```
set.seed(650)
mns<-NULL
for (i in 1 : 1000) mns = c(mns, mean(rexp(40,0.2)))
ggplot(data.frame(M=mns), aes(x=M)) + geom_histogram(aes(y=..density..),color="grey")+xlab("sample mean
```



then we can see that the distribution of the mean follows a normal distribution