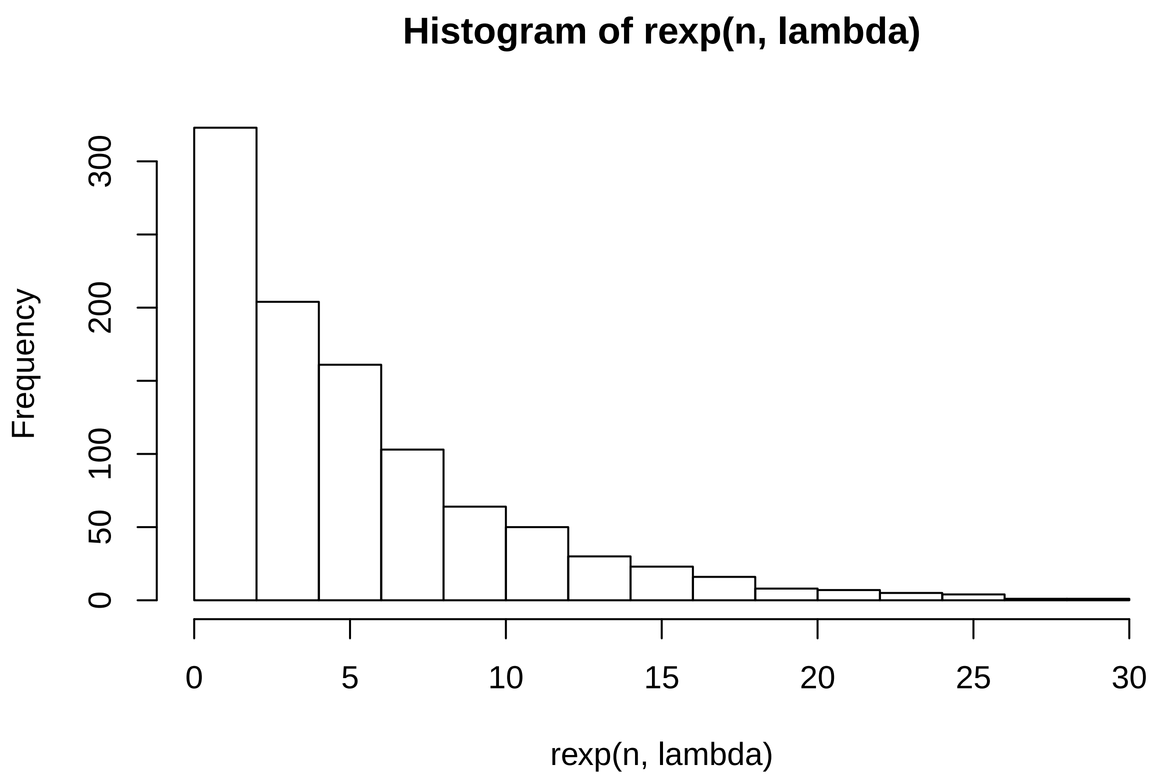
Statistical Inference: Course Project

1. Show where the distribution is centered at and compare it to the theoretical center of the distribution.

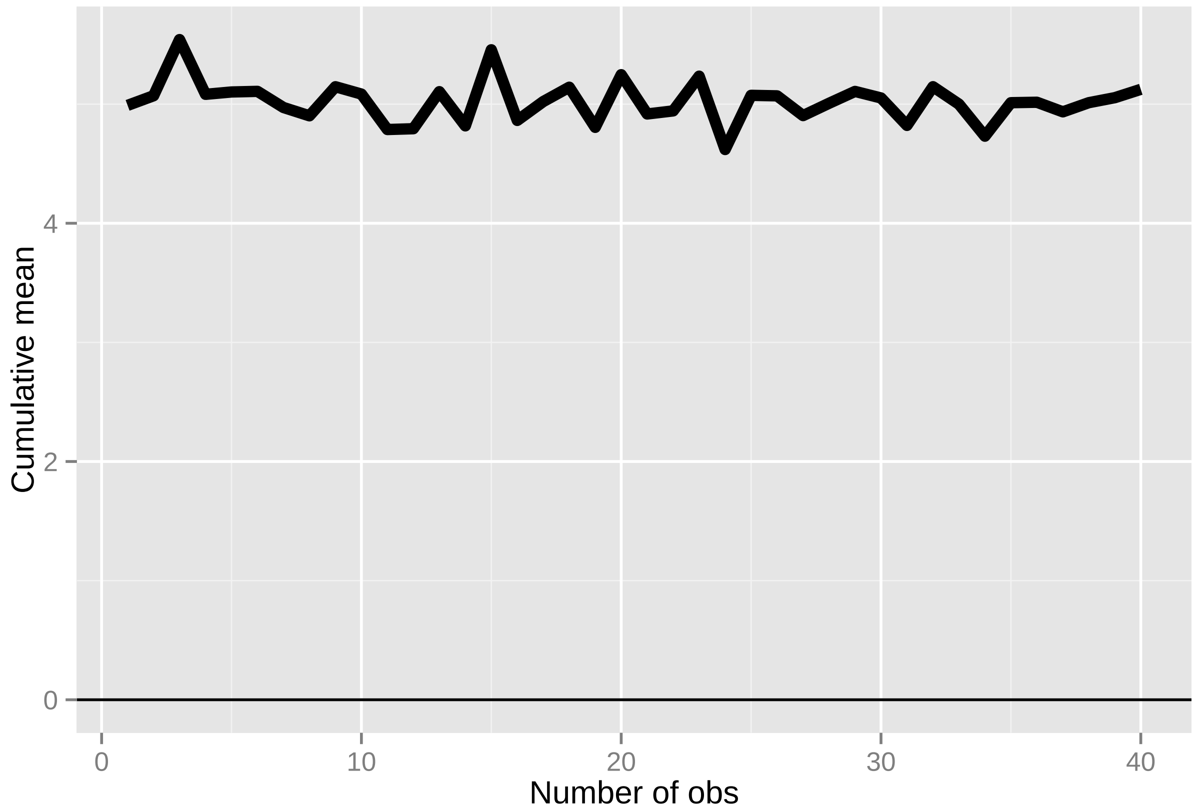
We can visualize the distribution with the histogram:

```
set.seed(5678)
n = 1000; lambda = 0.2;
hist(rexp(n, lambda))
```



n

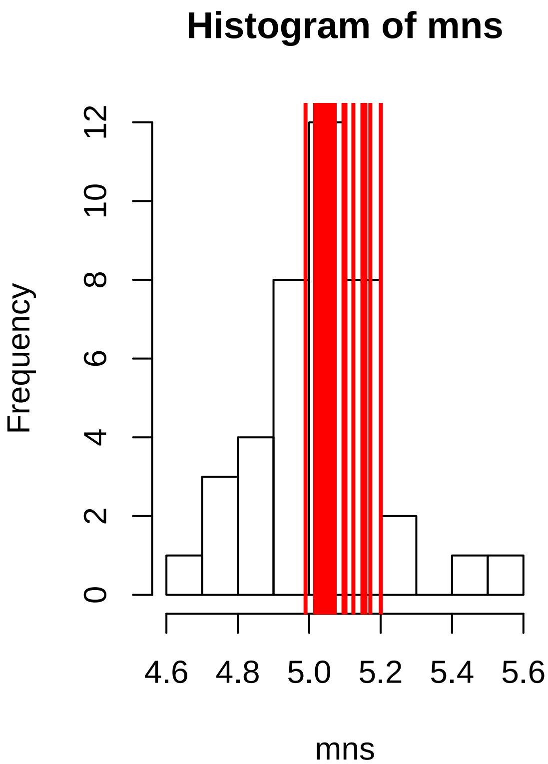
```
par(mfrow=c(1,2))
mns = NULL
for (i in 1 : 40) mns = c(mns, mean(rexp(n, lambda)))
hist(mns)
cltmean = cumsum(mns) / (1 : 40)
abline(v=cltmean, col="red", lwd=2)
g <- ggplot(data.frame(x = 1 : 40, y = mns), aes(x = x, y = y))
g <- g + geom_hline(yintercept = 0) + geom_line(size = 2)
g <- g + labs(x = "Number of obs", y = "Cumulative mean")
g
```

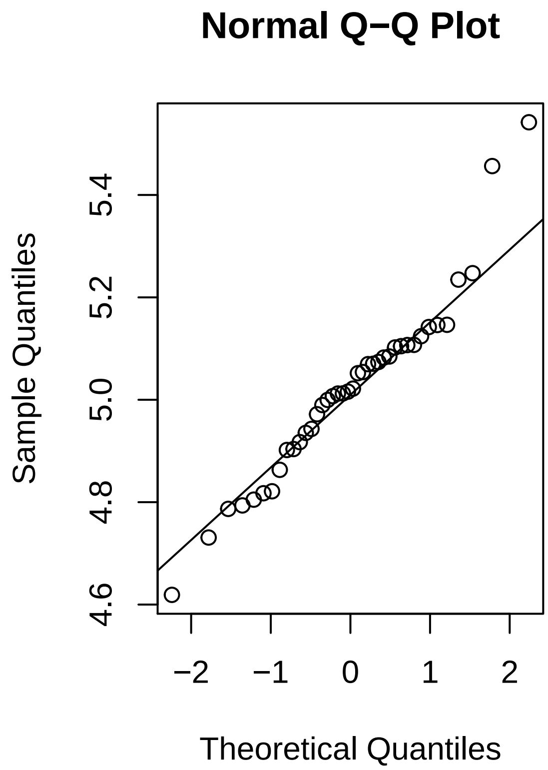


The expected value of an exponentially distributed random variable X with rate parameter λ is given by $E[X] = \frac{1}{\lambda} = \frac{1}{20} \approx 0.05$. We run a Monte Carlo simulation

we can visually see this is approximately normal:

```
par(mfrow=c(1,2))
hist(mns)
cltmean = cumsum(mns) / (1 : 40)
abline(v=cltmean, col="red", lwd=2)
qqnorm(mns); qqline(mns)
```





2.

Show how variable it is and compare it to the theoretical variance of the distribution.

The theretical variance of the distribution is $Var[X] = \frac{1}{\lambda^2} = \frac{1}{20^2} \approx 0.0025$. We run a Monte Carlo simulation

3. Show that the distribution is approximately normal.