Simulation of an Exponential random variable

Giacinto Maggiore

April 19, 2016

The goal of the project is to investigate a 0.2-exponential distribution in R and the distribution of avarages of 40 0.2-exponentials.We will make a comparision between the simulated distribution from a sample of 1000 means of 40 exponentials and the theoretical results of CLT.

1 Theoretical mean and sample mean

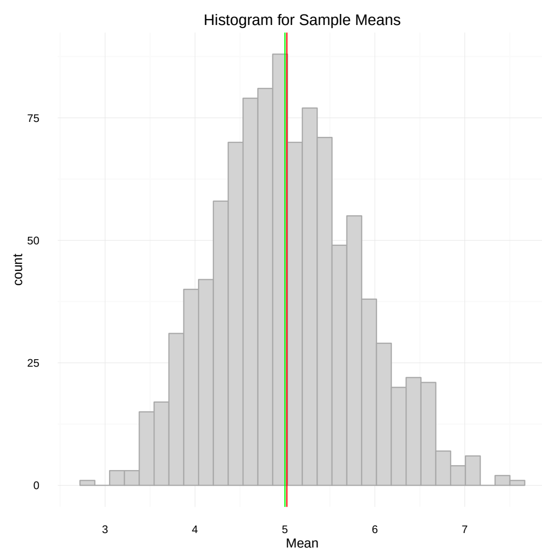
First of all we construct the sample means.

```
set.seed(20893)
lambda<-.2
n<-40
nsim<-1000
sample<-matrix(data = rexp(n*nsim,lambda),nrow = 40,
               ncol = 1000)
sampleMeans<-colMeans(sample)
m<-mean(sampleMeans)
```

The mean of the sample is 5.021, that is close to the theoretical mean $\frac{1}{\lambda} = 5$. This is underlined in the following plot:

```
qplot(sampleMeans,
         main = "Histogram for Sample Means",
         fill=I("lightgrey"),
         col=I("darkgrey"))+
         xlab("Mean")+
         ylab("count")+
         geom_vline(xintercept = mean(sampleMeans),col="red")+
         geom_vline(xintercept = 5,col = "green")+
         theme_minimal()

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



2 Theoretical variance and sample variance

In this section we will compare the variance of the sample with the theorical variance of the distribution.

```
sampleVariance <- (sd(sampleMeans))^2
realVariance <- 1/((lambda^2)*n)
```

The sample variance is 0.629 and the theoretical one is 0.625, so they are very close.

3 Distribution of the sample

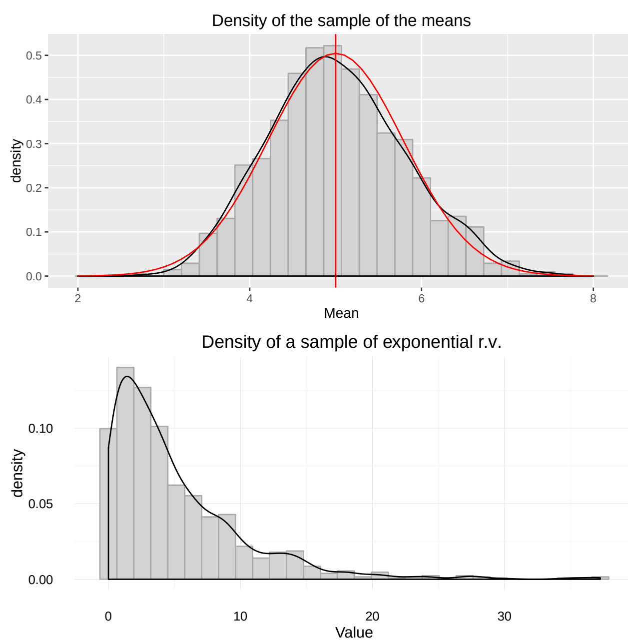
```
x <- seq(2,8,by=0.1)

g<-ggplot() +
geom_histogram(aes(x=sampleMeans,y=..density..),
               position="identity",fill=I("lightgrey"),
               col = ("darkgrey"))+
geom_density(aes(x=sampleMeans,y=..density..))+
ggtitle("Density of the sample of the means")+
xlab("Mean")+
geom_line(aes(x=x,
        y=dnorm(x,
                mean = 1/lambda,
```

```
                 sd = 1/(lambda*sqrt(40)))),
          color="red")+
geom_vline(xintercept = 5,col = "red")

sample<-rexp(nsim,lambda)
p<-ggplot()+
    geom_histogram(aes(x=sample,y=..density..),
              position="identity",fill=I("lightgrey"),
              col = ("darkgrey"))+
    geom_density(aes(x=sample,y=..density..))+
    ggtitle("Density of a sample of exponential r.v. ")+
    xlab("Value")+
    theme_minimal()
grid.arrange(g,p)

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



In the first plot the black line represents the simulated distribution density and it is very close to the normal distribution density (represented by red line) predicted by Central Limit Theorem. Observe that there is a lot of difference between the distribution of a large collection of averages of 40 exponentials and the distribution of a large collection of random exponentials. In fact if the first one is approximately normal, the second one has an exponential density, as can be seen in the second plot.