Simulation using the exponential distribution

Vaibhav Palve

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.2.4
```

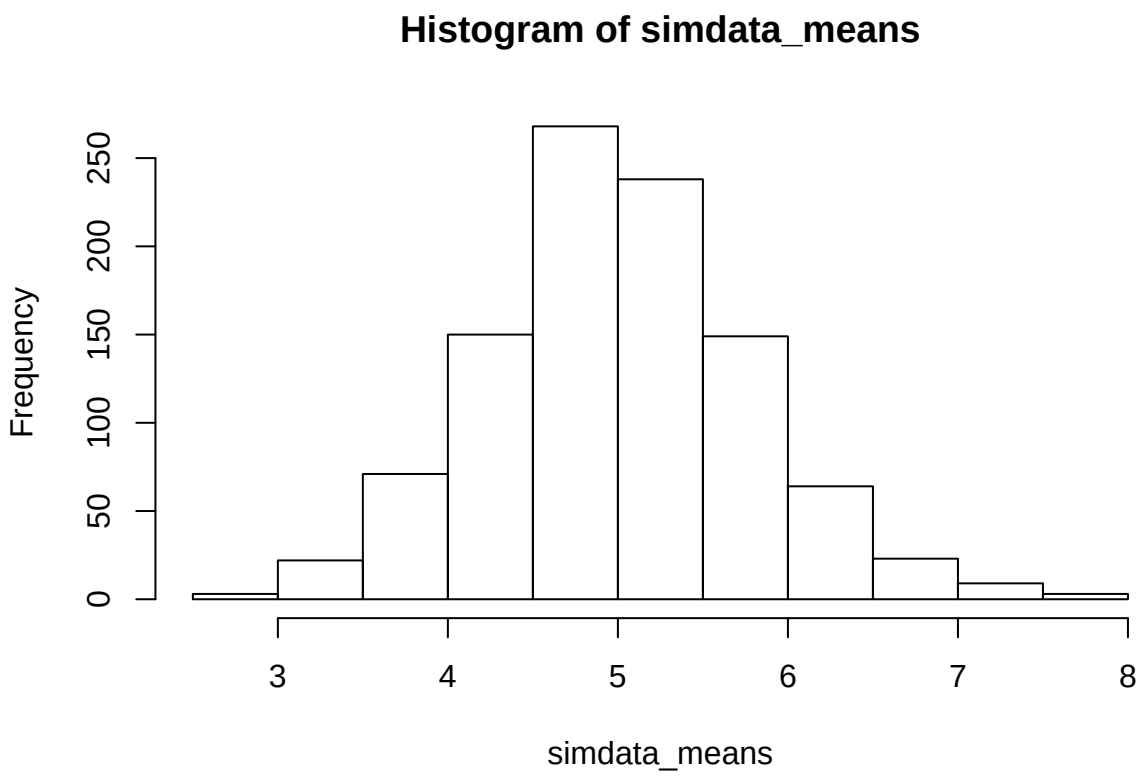
Now, set the variables as defined in the problem. number of values (n) = 40 lambda = 0.2 number of iterations, as least 1000. nosim=1000. Initialize the variables.

```
nosim <- 1000;
n <- 40;
lambda <- 0.2;
```

Using the rexp function, along with the matrix function develop a dataset with the mean and lambda specified above

Also, the theoretical mean is given as 1 / lambda or 1 / 0.2. The actual mean of the generated data can be calculated by using the apply & mean functions to obtain a mean for each row and then taking the mean of those numbers.

```
 ## values
simdata <- matrix(rexp(nosim * n, rate=lambda), nosim);
simdata_means <- apply(simdata, 1, mean);
hist(simdata_means);
```



```r
theoretical_mean <- 1/lambda;

print (paste("Theoretical center of the distribution = ",theoretical_mean));
```

```
## [1] "Theoretical center of the distribution =  5"
```

```r
print (paste("Actual center of the distribution based on the simulations = ",
             round(mean(simdata_means), 3)));
```

```
## [1] "Actual center of the distribution based on the simulations =  5.012"
```

```r
theoretical_var <- (1/lambda)^2/n;
theoretical_sd <- 1/lambda/sqrt(n);
print (paste("Theoretical variance = ", theoretical_var));
```

```
## [1] "Theoretical variance =  0.625"
```

```r
print (paste("Actual variance based on the simulations=",
             round(var(simdata_means), 3)));
```

```
## [1] "Actual variance based on the simulations= 0.617"
```

```r
print (paste("Theoretical standard deviation = ",
                     round(theoretical_sd, 3)));
```

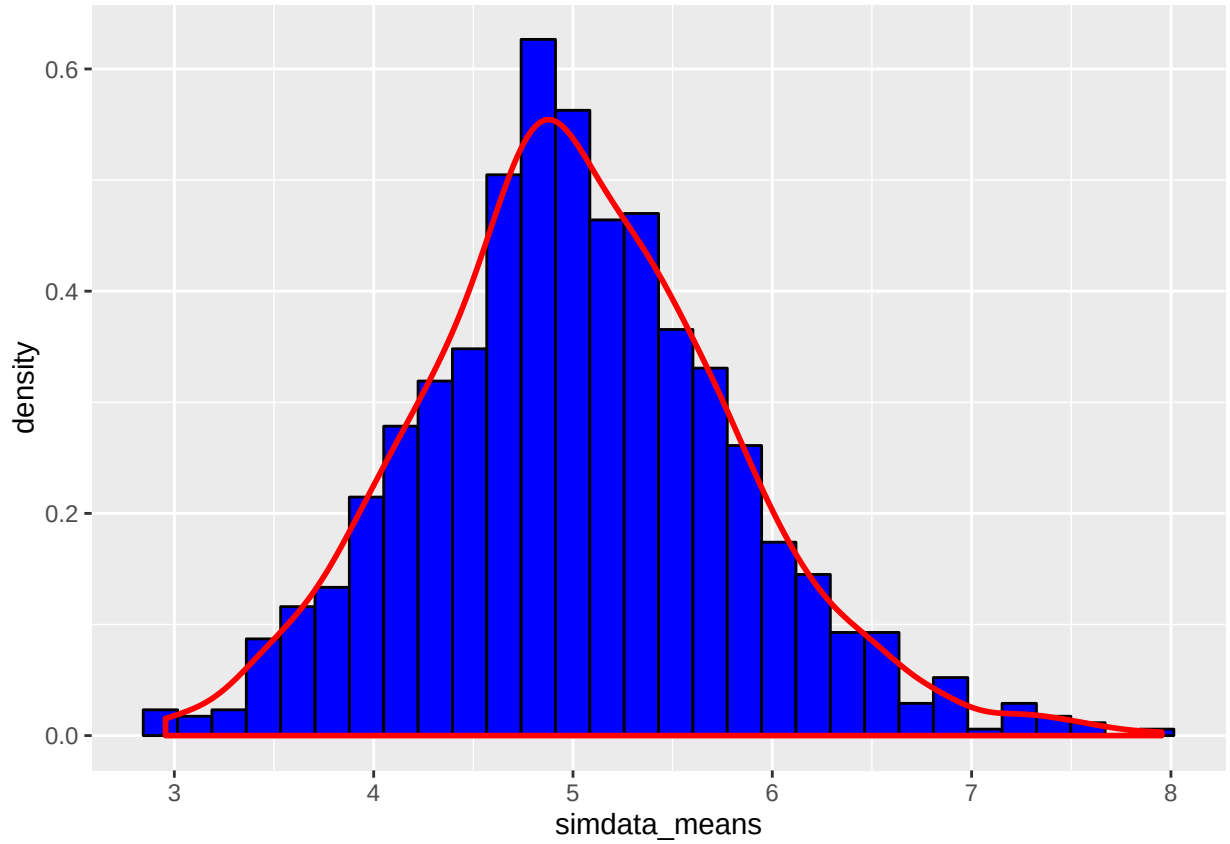
```
## [1] "Theoretical standard deviation =  0.791"
```

```r
print (paste("Actual standard deviation based on the simulations = ",
                     round(sd(simdata_means), 3)));
```

```
## [1] "Actual standard deviation based on the simulations =  0.786"
```

```r
plotdata <- data.frame(simdata_means);
m <- ggplot(plotdata, aes(x = simdata_means));
m <- m + geom_histogram(aes(y=..density..), colour="black",
                                fill = "blue")
m + geom_density(colour="red", size=1);
```

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



To answer question 1 - I have demonstrated the actual mean compared to the theroetical mean. You can check the value in the output above. To Answer question 2 - I have demonstrated the actual standard deviation and variance compared to their theoretical values. To answer question 3 - I graphed and was able to show that both the calculated curve is very close to the theoretical normal curve therefore proving the Central Limit Theory.