Inference Assignment 1

Moritz Schneider

Monday, December 15, 2014

```
library(ggplot2)
```

Assignment:

The exponential distribution can be simulated in R with rexp(n, lambda) where lambda is the rate parameter. The mean of exponential distribution is 1/lambda and the standard deviation is also also 1/lambda. Set lambda = 0.2 for all of the simulations. In this simulation, you will investigate the distribution of averages of 40 exponential(0.2)s. Note that you will need to do a thousand or so simulated averages of 40 exponential.

Illustrate via simulation and associated explanatory text the properties of the distribution of the mean of 40 exponential(0.2)s. You should

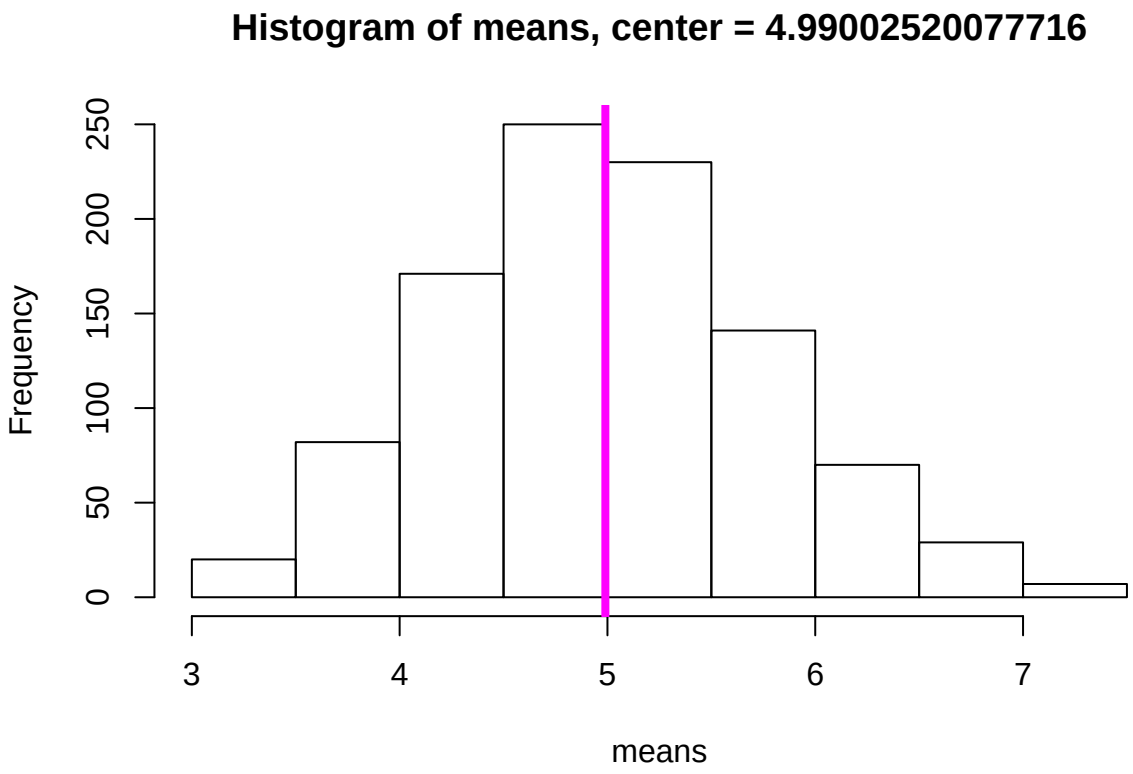
1. Show where the distribution is centered at and compare it to the theoretical center of the distribution.
2. Show how variable it is and compare it to the theoretical variance of the distribution.
3. Show that the distribution is approximately normal.

Part 1

I will use set.seed(1) as seed for this and all following parts.

According to the central limit theorem the distribution of sample means should be centered at the population mean, i.e. 5. To confirm this, we calculate the mean of 40 exponentials with lambda = 0.2, repeat this 1000 times and plot the resulting distribution.

```
set.seed(1)
lambda <- 0.2
nm <- 40
reps <- 1000
means = NULL
for (i in 1:reps){
        means = c(means, mean(rexp(nm, lambda)))
}
sim.mean <- mean(means)
hist(means,
     main = paste("Histogram of means, center = ",sim.mean, sep=""))
abline( v = sim.mean, col = "magenta", lwd =4)
```



The magenta colored line gives us the mean of the distribution, which is 4.99 and therefore a very good estimate of the population mean.

Part 2

Let's calculate the variance of the distribution we just sampled:

```
var(means)
```

```
## [1] 0.6111165
```

The theoretical value for the variance should be:

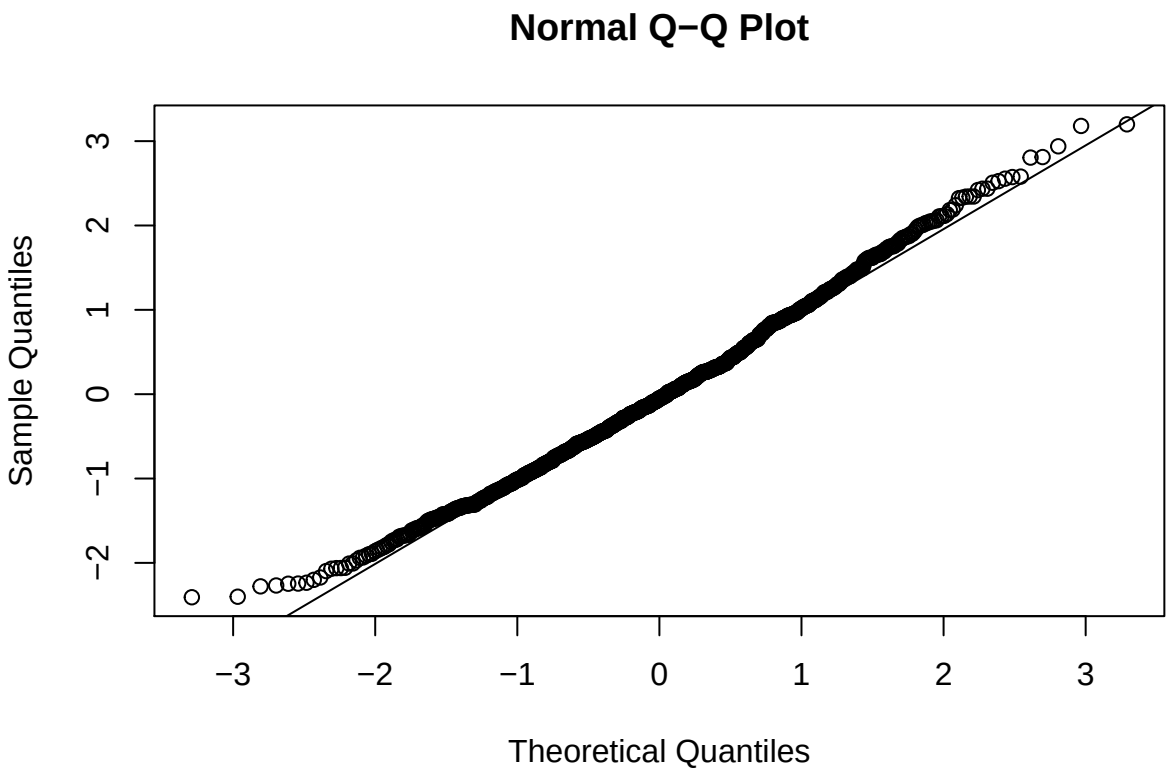
```
1/(lambda^2*nm)
```

```
## [1] 0.625
```

Again, the sampled variance is very close to the theoretical value.

Part 3

To show that the sampled distribution is approx. normal we transform our distribution to a distribution with mean = 0 and sd = 1 and then take a look at the Normal Q-Q Plot of our sampled distribution



Here we can see, that our distribution is very close to normal, which is expected as the distribution of averages tends to approximate a normal distribution as the sample size used to calculate the mean gets larger.