Statistical Inference Course Project

Rui Mendes

24-01-2015

Part1: "Simulation Exercises"

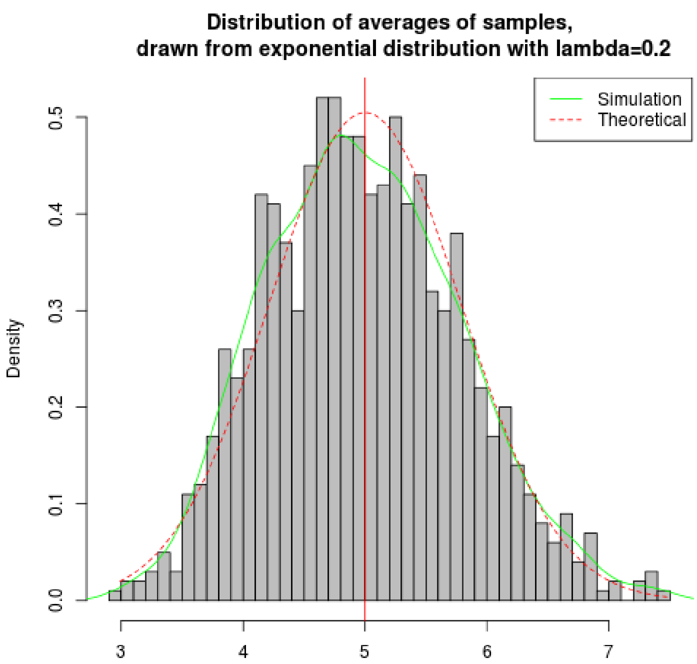
The exponential distribution can be simulated in R with `rexp(n, lambda)` where `lambda` λ is the rate parameter. The mean of exponential distribution is $1/\lambda$ and the standard deviation is also $1/\lambda$. In this project, we set $\lambda = 0.2$ for all simulations.

In this simulation, we will investigate the distribution of averages of 40 exponentials.

Let's do a thousand simulated averages of 40 exponentials.

```
## Define seed (to make reproducible)
set.seed(3)

## Set the value of lambda
lambda <- 0.2

## Setnumber of simulations
num_sim <- 1000

## Set the sample size
sample_size <- 40

## Define a matrix with the sample
sim <- matrix(rexp(num_sim*sample_size, rate=lambda), num_sim, sample_size)
row_means <- rowMeans(sim)
```

The distribution of sample means is as follows.

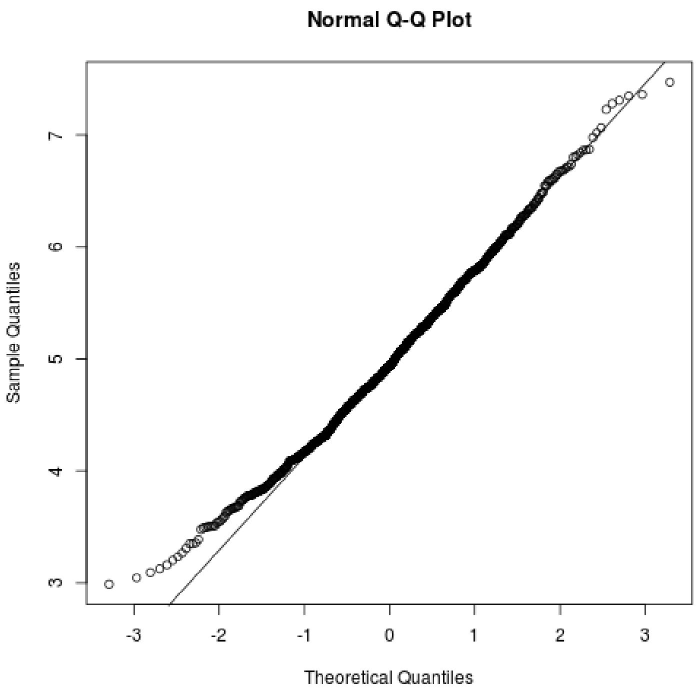


The distribution of sample means is centered at 4.9866197 and the theoretical center of the distribution is λ^{-1} = 5.

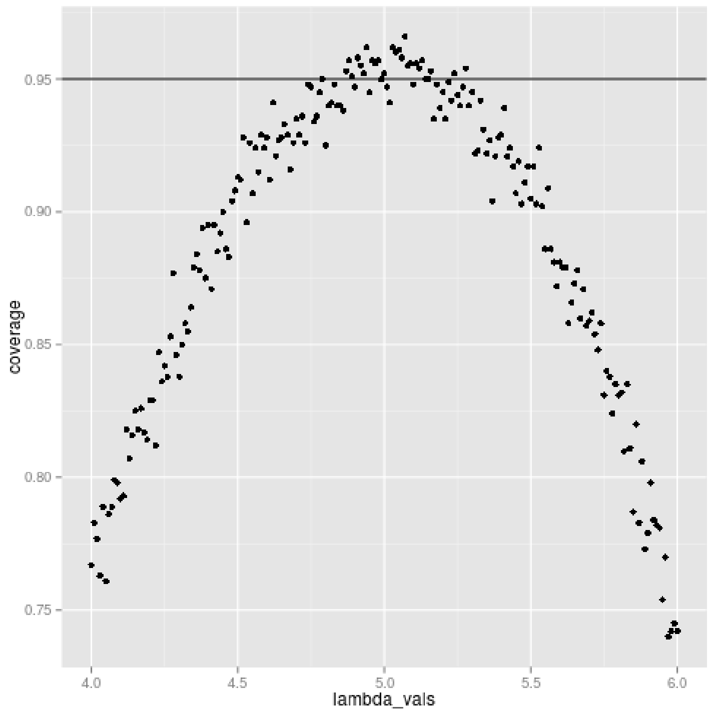
The variance of sample means is 0.6257575 where the theoretical variance of the distribution is $\sigma^2/n = 1/(\lambda^2 n) = 1/(0.04 \times 40)$ = 0.625.

Due of the central limit theorem, the averages of samples follow the normal distribution.

The figure plotted above also shows the density computed using the histogram and the normal density plotted with theoretical mean and variance values. Also, the q-q plot below suggests the normal distribution



Finally, let’s see the evaluation of the coverage for the confidence interval with $1/\lambda = \bar{X} \pm 1.96\,\frac{S}{\sqrt{n}}$



The 95% confidence intervals for the rate parameter (λ) to be estimated ($\hat{\lambda}$) are $\hat{\lambda}_{low} = \hat{\lambda}(1 - \frac{1.96}{\sqrt{n}})$ agnd $\hat{\lambda}_{upp} = \hat{\lambda}(1 + \frac{1.96}{\sqrt{n}})$. As can be seen from the plot above, for selection of $\hat{\lambda}$ around 5, the average of the sample mean falls within the confidence interval at least 95% of the time. Note that the true rate, λ is 5.

This file can be found here (https://github.com/rui-mendes/StatisticalInference/tree/master/Part1) and/or here (http://rpubs.com/ruimendes/statistical-inference-proj-part1).