Statistical Inference Coursera Assignment: Part I

ROSTISLAV UHLIR

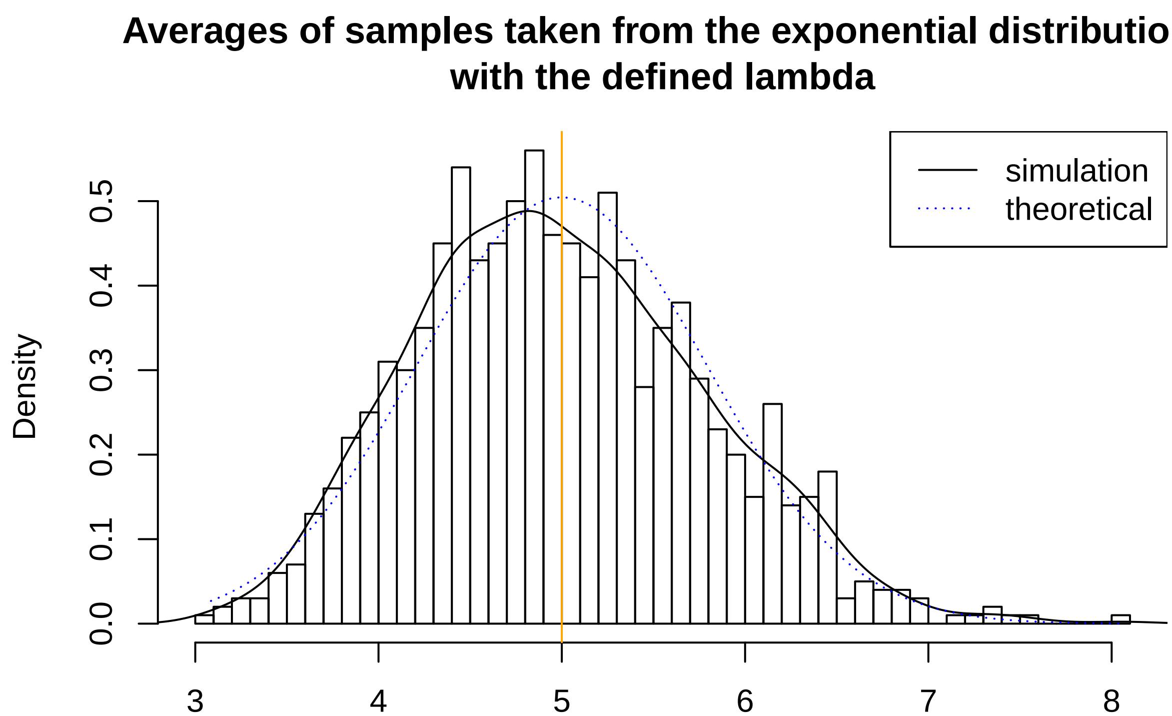
22 September 2015

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 $1/\lambda$. For this simulation, we set $\lambda = 0.2$. In this simulation, we investigate the distribution of averages of 40 numbers sampled from exponential distribution with $\lambda = 0.2$.

Thousand simulated averages of 40 exponentials:

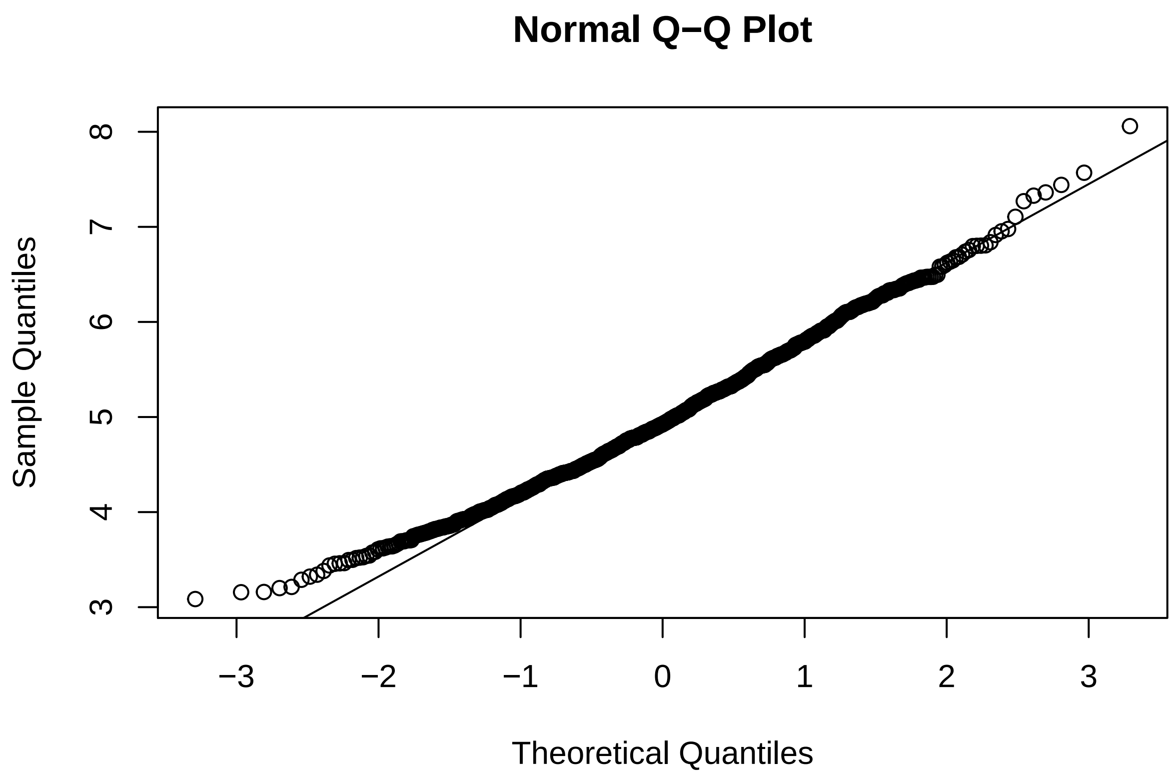
```
set.seed(1)
lambda <- 0.2
num_sim <- 1000
sample_size <- 40
sim <- matrix(rexp(num_sim*sample_size, rate=lambda), num_sim, sample_size)
row_means <- rowMeans(sim)
```

Distribution of sample means:

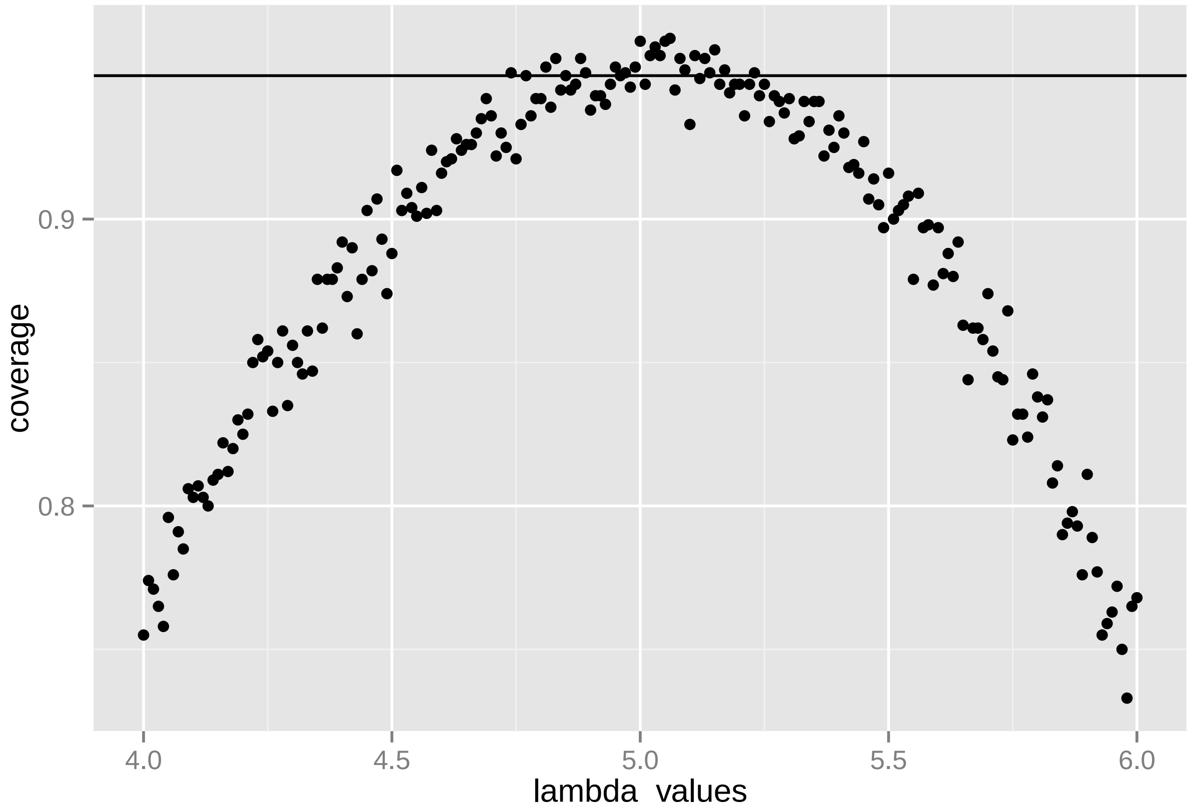


The distribution of sample means is centered at 4.9900252 and the theoretical centre of the distribution is $\lambda^{-1} = 5$. The variance of sample means is 0.6177072 where the theoretical variance of the distribution is $\sigma^2/n = 1/(\lambda^2 n) = 1/(0.04 \times 40) = 0.625$.

The averages of samples follow normal distribution, this is due to the fact that the central limit theorem applies. The graph above shows the density calculated using the histogram and the normal density taking into account the theoretical mean and variance values. The q-q plot helps to show the normality of the distribution.



Evaluation of the coverage of the confidence interval for $1/\lambda = \bar{X} \pm 1.96\frac{S}{\sqrt{n}}$



The 95% confidence intervals for the rate parameters (λ) to be estimated ($\hat{\lambda}$) are $\hat{\lambda}_{low} = \hat{\lambda}(1 - \frac{1.96}{\sqrt{n}})$ and $\hat{\lambda}_{upp} = \hat{\lambda}(1 + \frac{1.96}{\sqrt{n}})$. We observe from the graph plotted above that the selection of $\hat{\lambda}$ is around 5 and that the average of the sample mean is within the confidence interval at least 95% of the time. The true rate, λ is 5.