Simulation Exercise - Exponential Distribution

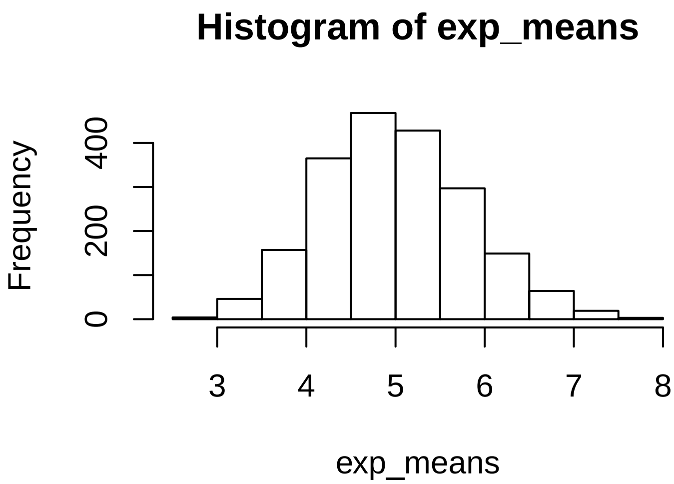
Ivan Filho

Introduction

I have created a GitHub repository where I've put the files related to this homework. It can be accessed via the following url: https://github.com/ivanmantova/coursera-statsinference.

Analysis

We are going to simulate 40 iid random variables drawn from an Exponential distribution with parameter $\lambda = .2$ and analyse the sampling distribution of their averages.



We have taken 2000 means from 40 iid draws from the distribution and plotted a histogram. According to the **Central Limit Theorem**, we know that the distribution of sample averages should look like a Gaussian (normal) distribution as the sample size increases. Also, according to the **Law of Large Numbers** the sample mean of iid samples is consistent for the population mean, as well as the sample variance and sample standard deviation.

Here we have a sample size of 40, which should be enough for our purposes.

We can see from the plot that the distribution does look aproximately normal. We are now going to perform some basic inference on these simulated samples.

Estimating the true (population) mean μ with $\bar{X}$

The mean of the Exponential distribution is $E[X] = \frac{1}{\lambda} = \lambda^{-1}$, so the mean of the population distribution with rate parameter $\lambda = 0.2$ is equal to 5 ($.2^{-1} = 5$).

Our estimate of the true mean of the distribution is the empirical mean $\bar{X}$, which is equal to:

```
## [1] 5.001
```

Variance and standard deviation

The variance of an exponentially distributed random variable X is given by

$$Var[X] = \frac{1}{\lambda^2} = \lambda^{-2}$$

which we can calculate since we are doing simulations and we know the true parameter λ. So in our case $\lambda^{-2} = .2^{-2} = 25$. The standard deviation is given by the square root of the variance $\sqrt{\lambda^{-2}} = \sqrt{25} = 5$.

Standard error of the mean

We can also calculate the standard error of the sampling distribution of the mean by

$$\frac{\sqrt{\lambda^{-2}}}{\sqrt{n}} = \frac{\sqrt{25}}{\sqrt{40}} = .7905694$$

which we can estimate with the standard deviation of our sample means

$$SD_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

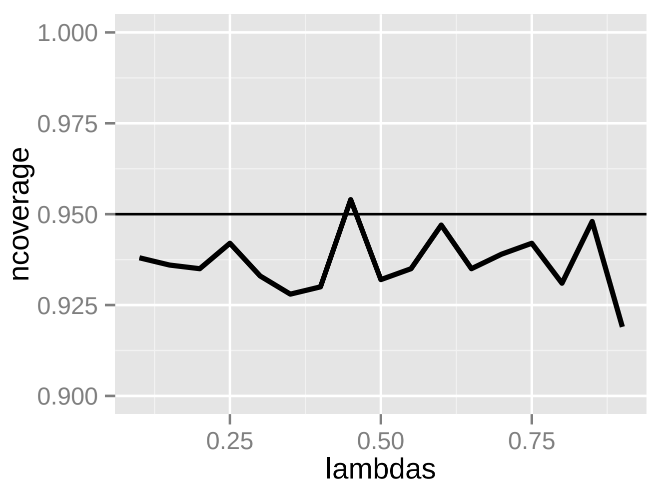
```
## [1] 0.8214
```

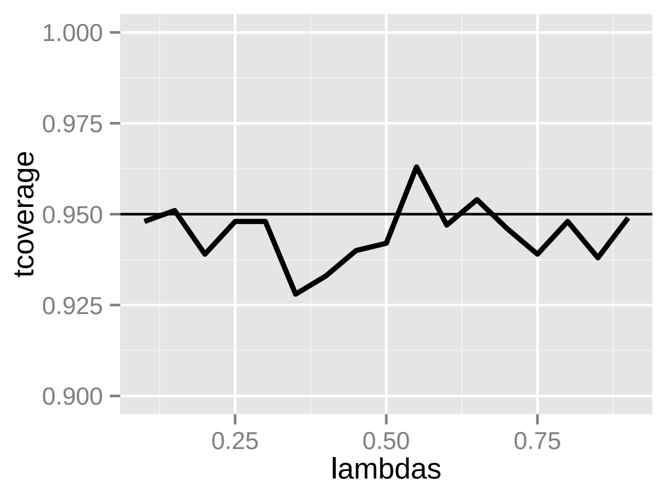
Confidence Interval for the mean

Given that we know the properties stated at the beginning and now that we have our estimates $\bar{X}$ and S we can calculate a 95% confidence interval for the mean with the formula $\bar{X}_n \pm 1.96\sigma/\sqrt{n}$:

```
## [1] 4.746 5.255
```

Let's evaluate the coverage of the 95% confidence interval for $\frac{1}{\lambda}$ for various λ ranging from .1 to .9. I also want to compare the coverage of the 95% Confidence Interval for Normal Distribution and the coverage of 95% Confidence Interval for t-Distribution:





As we can see, because of the small n (40), the t-distribution confidence interval does a slighly better job at covering the true value of the mean μ.

```
## Average ncoverage = 0.936706
```

```
## Average tcoverage = 0.944765
```