Statistical Inferential Analysis Course Work

Summary

This a presentation of course work for Coursera Inferential Statistics online course. In this study distribution of 1000 mean vlues for sets containg 40 random exponential distributions samples each. Results confirm that means are normally distributed with their mean aproximating the theoretical mean and variance approaching that of poplation divided by the sample size. Calculation done using R 3.1.1 language on Windows 8 64bit platform

Setup

```
sim_number <- 1000
n <- 40
lambda = 0.2

set.seed(511)
x <- replicate(sim_number, mean(rexp(40, rate=lambda)))
```

Results

1. Show where the distribution is centered at and compare it to the theoretical center of the distribution. Expected sample mean

$$\bar{X} = \sigma = \frac{1}{\lambda}$$

$1/0.2 =$ `5`

Calculated simulated sample mean

```
smean <- sum(x) / sim_number
smean
```

```
## [1] 4.967
```

2. Show how variable it is and compare it to the theoretical variance of the distribution. Expected value for sample variance

$$s^2 = \frac{\sigma^2}{n} = \frac{(\frac{1}{\lambda})^2}{n} = \frac{1}{\lambda^2 n}$$

$1/(0.2^2 * 40) =$ `0.625`

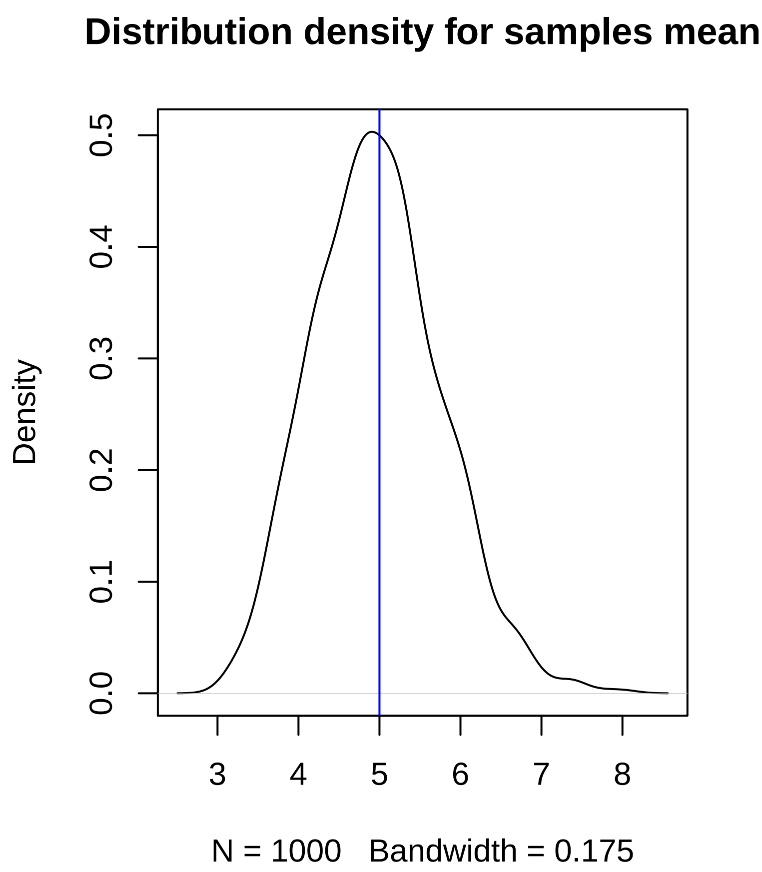
Calculated sample variance: $s^2_{n-1} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} =$

```
sum((x - smean)^2)/(sim_number-1)
```

```
## [1] 0.621
```

3. Show that the distribution is approximately normal. Distribution density function looks approximately like a 'bell' shape:

```
plot(density(x),main="Distribution density for samples mean")
abline(v=1/lambda, col="blue")
```



4. Evaluate the coverage of the confidence interval for $\frac{1}{\lambda}$:

$$\bar{X} \pm 1.96\frac{S}{\sqrt{n}}$$

To answer this question we just count number of sampled means out of `1000` simulations in CI, c.e. whose distansce from the population mean `5` is in boundaries of the fomula value

```
sum(abs(x - 1/lambda ) < 1.96 * smean / sqrt(n))
```

```
## [1] 951
```

which makes `95.1`% of `1000` simulations.