StatInference_Courseproject1

Frank Fichtenmueller

3 November 2016

Statistical Inference Course Project 1

Overview

The scope of this final Course Project is to investigate the exponential distribution in R from the perspective of the Central Limit Theorem. From the Theorem follows, that a randomized redraw of sample distributions will construct a normalized distributiof over its mean. To simulate an exponential distribution in R, i will make use of the rexp(n, lambda) function, where lambda is the rate parameter. I will set lambda = 0.2 for all of the simulations and document the distribution of averages of 40 different exponential settings. The sample size is set to be 1000 draws per simulation.

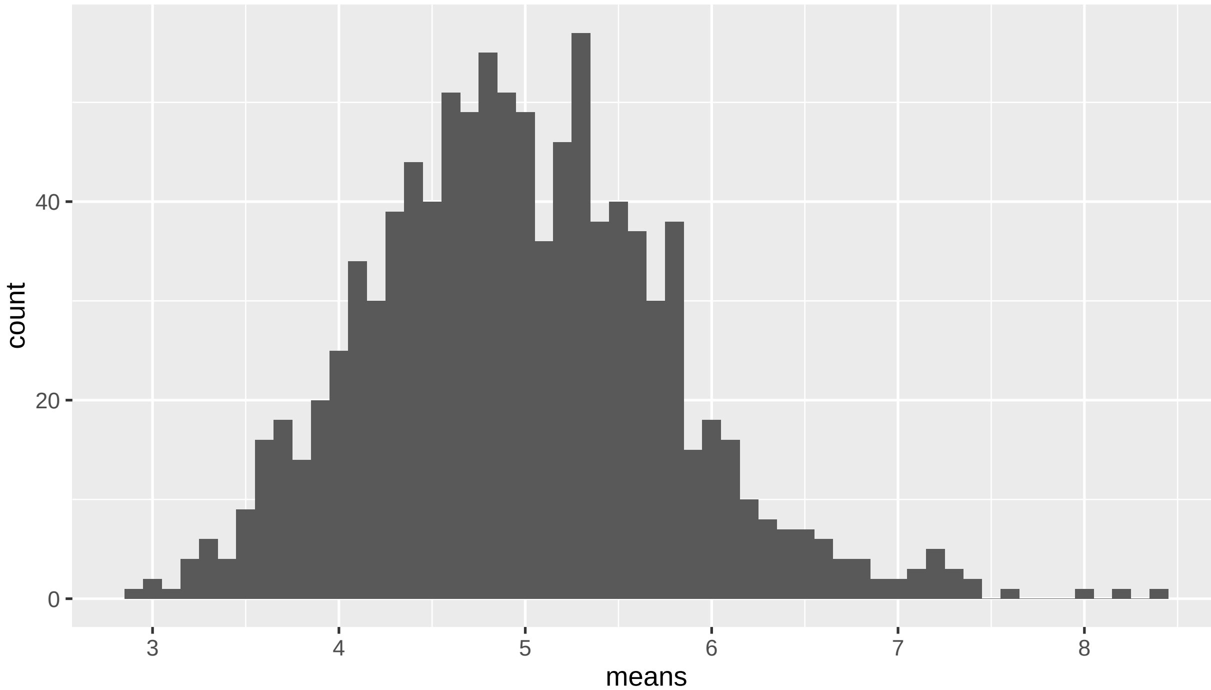
Simulations

```r
# load neccesary libraries
library(ggplot2)

# set constants
lambda <- 0.2 # lambda for rexp
n <- 40 # number of exponetials
nSumim <- 1000 # number of tests

# set the seed
set.seed(45876235)

# run the test resulting in n x nSumim matrix
expDistr <- matrix(data=rexp(n * nSumim, lambda), nrow=nSumim)
expDistrMean <- data.frame(means=apply(expDistr, 1, mean))
```



Sample Mean versus Theoretical Mean

The expected mean μ of a exponential distribution of rate λ is

$\mu = \frac{1}{\lambda}$

```
mu <- 1/lambda
mu
```

```
## [1] 5
```

Let $\bar{X}$ be the average sample mean of 1000 simulations of 40 randomly sampled exponential distributions.

```
meanOfMeans <- mean(expDistrMean$means)
meanOfMeans
```

```
## [1] 4.969255
```

As you can see the expected mean and the avarage sample mean are very close

Sample variance versus Theoretical variance

The expected standard deviation σ of a exponential distribution of rate λ is

$\sigma = \frac{1/\lambda}{\sqrt{n}}$

The e

```
sd <- 1/lambda/sqrt(n)
sd
```

```
## [1] 0.7905694
```

The variance var of standard deviation σ is

$var = \sigma^2$

```
var <- sd^2
var
```

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

Let var_x be the variance of the average sample mean of 1000 simulations of 40 randomly sampled exponential distribution, and σ_x the corresponding standard deviation.

```
sd_x <- sd(expDistrMean$means)
sd_x
```

```
## [1] 0.8093777
```

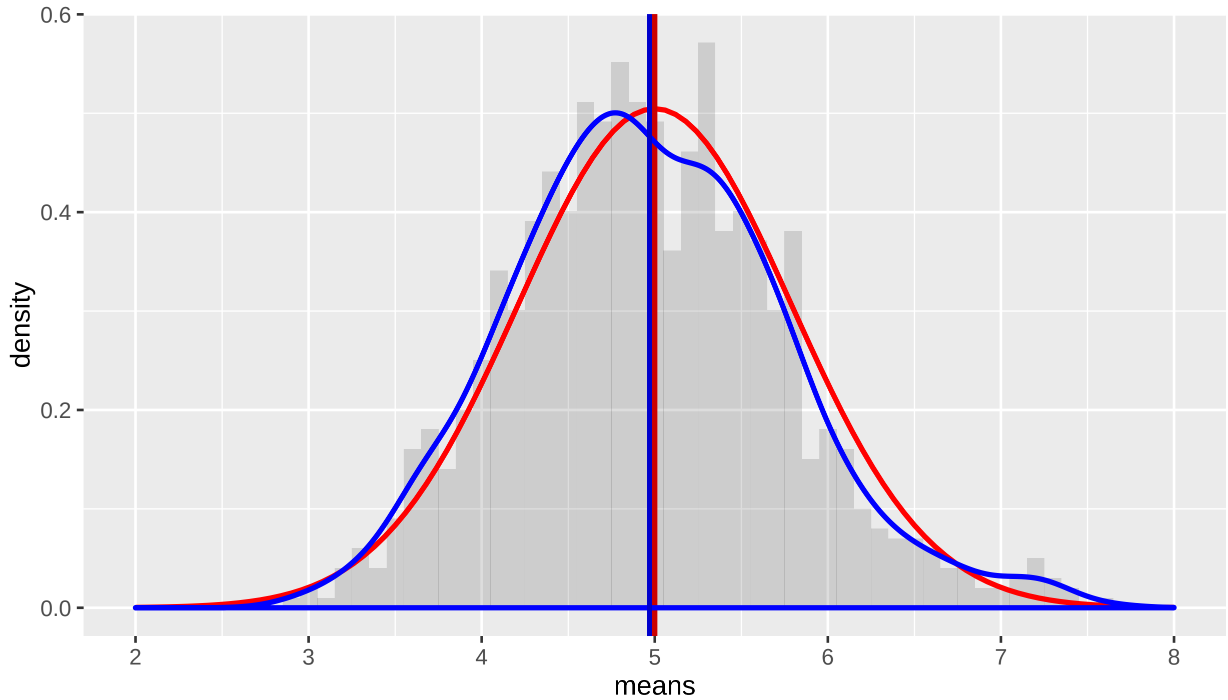
```
var_x <- var(expDistrMean$means)
var_x
```

```
## [1] 0.6550922
```

As you can see the standard deviations are very close Since variance is the square of the standard deviations, minor differnces will we enhanced, but are still pretty close.

Distribution

Comparing the population means & standard deviation with a normal distribution of the expected values.



As you can see from the graph, the calculated distribution of means of random sampled exponantial distributions, overlaps quite nice with the normal distribution with the expected values based on the given lamba