Statistical Inference Project

AVA

Thursday, January 22, 2015

Part A

Overview

In this project, we have investigated the exponential distribution in R and then, we have compared it with the Central Limit Theorem. The exponential distribution can be simulated in R with << rexp(n, lambda) >> where lambda is the rate parameter.

The mean of the exponential distribution is 1/lambda and the standard deviation is also 1/lambda. Lambda parameter was set to 0.2 for all simulations. We have investigated the distribution of averages of 40 exponentials.

We embeded the following R code chunks :

Simulations

The distribution of averages

1. We have considered the distribution of a large collection of averages of 40 exponentials (the distribution of 1000 averages of 40 random exponentials).

```
lambda <-  0.2
n <- 40
nosim <- 1000
set.seed(1)
collectiondata <- rexp(nosim * n, lambda)
matrixdata <- matrix(collectiondata, nrow=nosim, ncol=n)
str(matrixdata)
```

```
##  num [1:1000, 1:40] 3.776 5.908 0.729 0.699 2.18 ...
```

```
means <-  apply(matrixdata, 1, mean)
str(means)
```

```
##  num [1:1000] 4.9 5.23 6.4 4.74 5.18 ...
```

```
# 1. Show the sample mean and compare it to the theoretical mean of the distribution.
mean(means)
```

```
## [1] 4.990025
```

```
1/lambda
```

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

```r
# 2. Show how variable the sample is (via variance) and compare it to the theoretical variance of the d
var(means)
```

```
## [1] 0.6177072
```

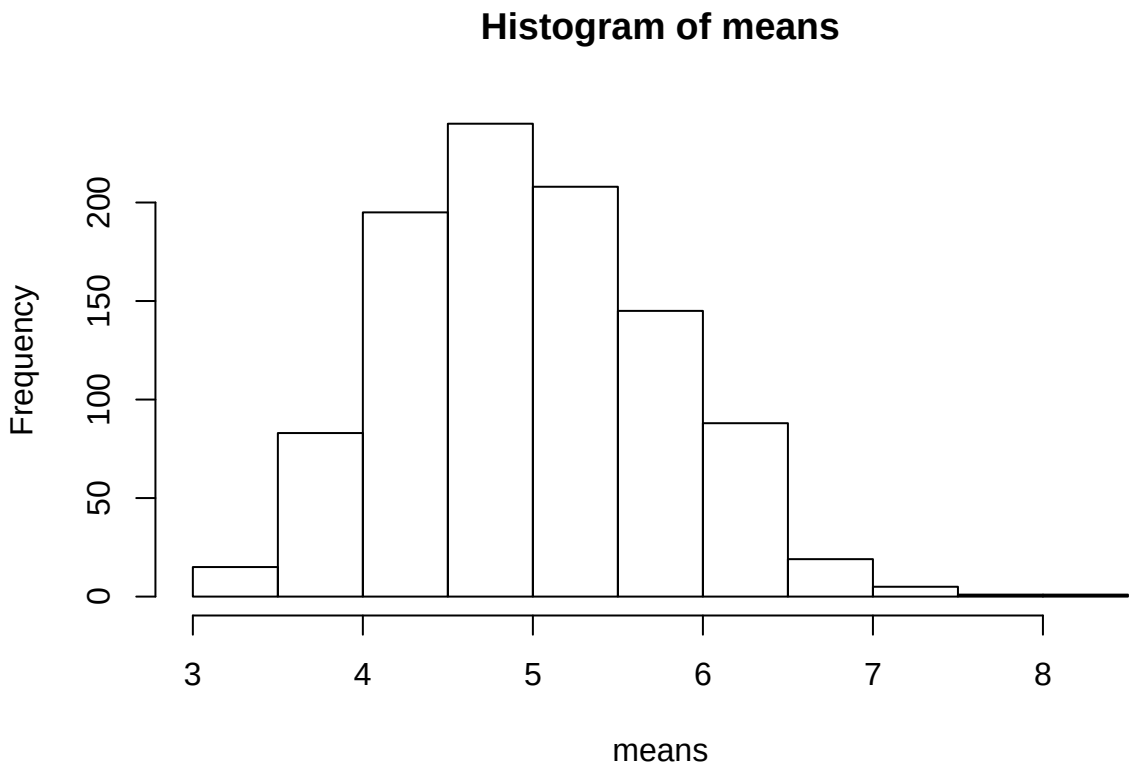
```r
sd(means)
```

```
## [1] 0.7859435
```

```r
(1/lambda)/sqrt(n)
```

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

```r
# 3. Show that the distribution is approximately normal.
hist(means)
```



The exponential distribution

2. Then we compared it with the distribution of a large collection of random exponentials (the distribution of 1000 random exponentials).

```r
lambda <-  0.2
n <- 40
set.seed(1)
expdistribution <- rexp(1000, lambda)
str(expdistribution)
```

```
##  num [1:1000] 3.776 5.908 0.729 0.699 2.18 ...
```

```r
mean(expdistribution)
```

```
## [1] 5.156513
```

```r
var(expdistribution)
```

```
## [1] 24.46583
```

```r
sd(expdistribution)
```

```
## [1] 4.946295
```

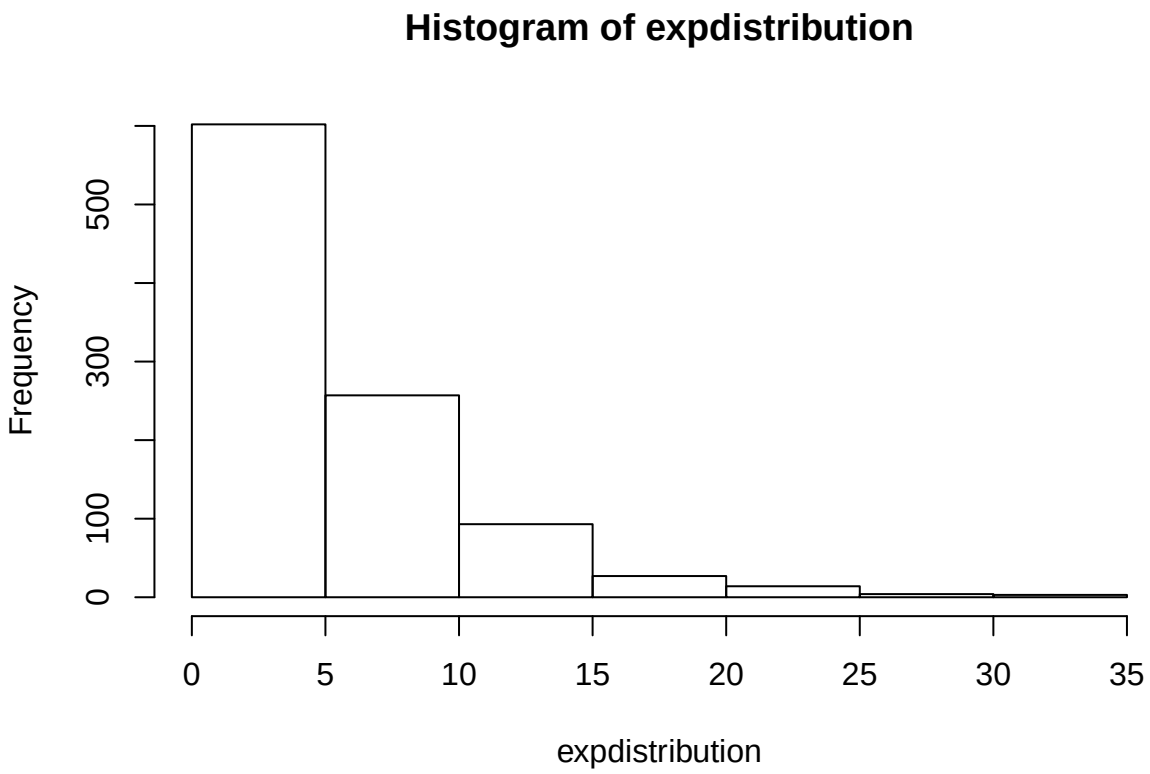
```r
1/lambda
```

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

```r
(1/lambda)/sqrt(n)
```

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

```r
hist(expdistribution)
```



Note that the mean of the distribution of 1000 averages of 40 random exponentials(0.2) is 4.99, which is very close to the expected mean (1/lambda = 5.0).