Part 1 - Statistical Inference Project

Tamer Salem

5/4/2020

Simulation Exercise

Overview

Investigate the exponential distribution in R and compare it w/the Central Limit Theorem.

As mandated by assignment, exponential distribution will be simulated in R using rexp(n, lambda) where lambda is the rate parameter. Distribution mean and standard deviation are 1/lambda and lambda = 0.2 for all simulations.

Assigment mandates investigation of "distribution of averages of 40 exponentials" w/1000 simulations.

Simulations

Initialization.

```
set.seed(54321)
lambda <- 0.2
n <- 40
simulationLimit <- 1000
standardDeviation <- 1/lambda
```

Sample Mean vs Theoretical Mean

```
simulation <- rexp(simulationLimit * n, rate = lambda)
simulatedMean <- mean(simulation)
simulatedMean
```

```
## [1] 4.990132
```

```
theoreticalMean <- 1/lambda
theoreticalMean
```

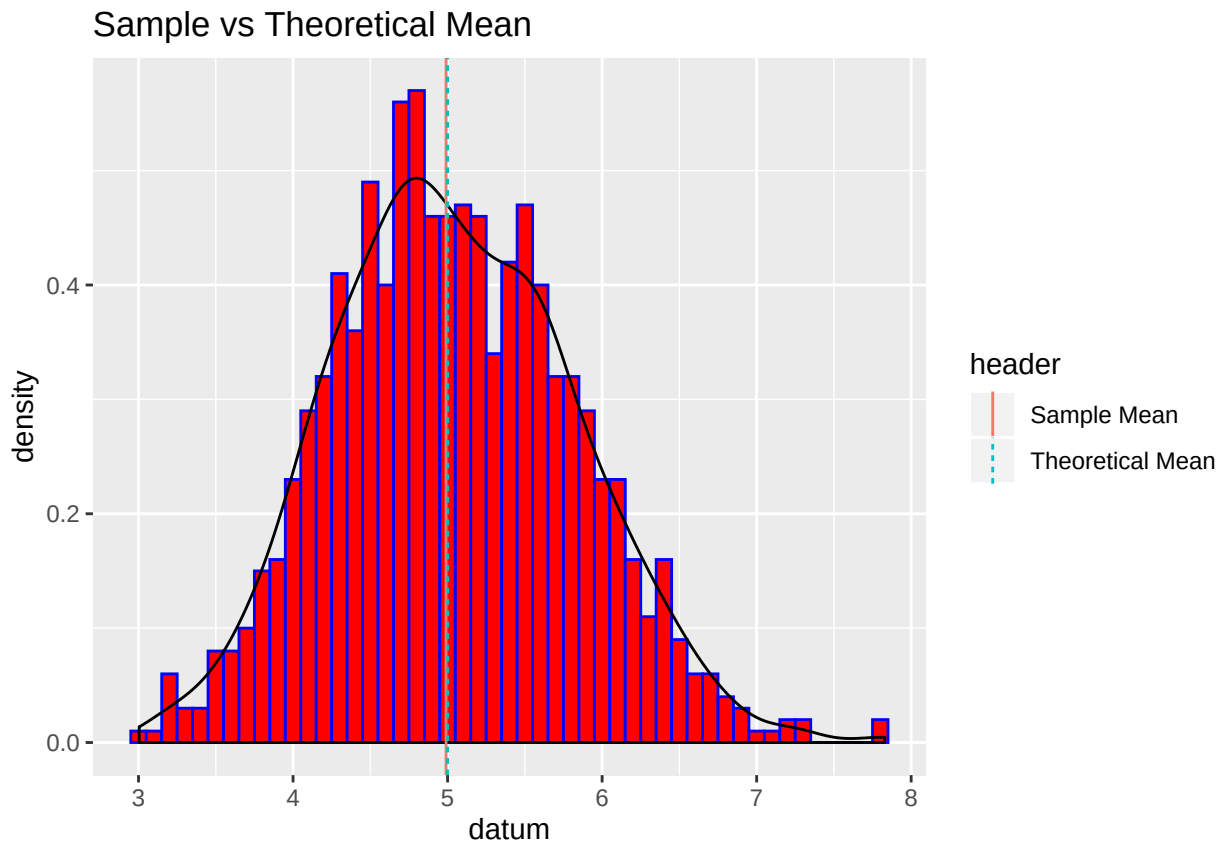
```
## [1] 5
```

Sample Variance vs Theoretical Variance

```
datum <- NULL
for (ii in 1:simulationLimit) {
  datum <- c(datum, mean(rexp(n, lambda)))
}
```

Illustrate sample mean vs theoretical mean.

```
library(ggplot2)
dataFrame1 <- data.frame(header = c("Sample Mean", "Theoretical Mean"), values = c(simulatedMean, theor
gg <- ggplot(NULL, aes(x=datum))
plot1 <- gg + geom_histogram(aes(y=..density..), color="blue", fill="red", binwidth=0.1) + labs(title =
plot1
```



Standard deviation

```
sd(datum)
```

```
## [1] 0.7773352
```

Variance

```
varianceDatum <- sd(datum)^2
varianceDatum
```

```
## [1] 0.60425
```

Theoretical variance

```
theoreticalDatum <- (1/lambda)^2/n
theoreticalDatum
```

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

Normal Distribution

The distribution plot (above) clearly exhibits gaussian bell curve shape, indicated normal distribution. Plot below uses qqnorm() to compare simulation to normal, indicating close approximation.

```
qqnorm(datum)
qqline(datum, col=10)
```