Statistical Inference - Project Part 1

Kari Innes
September 15, 2014

Simulated 40 exponential(0.2)s with lambda = 0.2 and 1000 simulations

Set variables

```
lambda = 0.2
n = 40
simulations = 1000

simulatedData <- matrix(0,nrow=simulations,ncol=n)

testResults <- data.frame(mean = numeric(0), sd = numeric(0), lowerCI = numeric(0),
upperCI = numeric(0), coverageTest = numeric(0))
```

1000 simulations

```
for (i in 1:simulations)
{
simulatedData[i,] <- rexp(n,lambda)
testResults[i,1] <- mean(simulatedData[i,])
testResults[i,2] <- sd(simulatedData[i, ])
testResults[i,3] <- testResults$mean[i] - 1 * qnorm(.975) * testResults$sd[i] / sqrt(n)
testResults[i,4] <- testResults$mean[i] + 1 * qnorm(.975) * testResults$sd[i] / sqrt(n)
testResults[i,5] <- testResults$lowerCI[i] < (1/lambda) & (1/lambda) < testResults$upperCI[i]


}
```

The expected value of an exponentially distributed random with rate parameter λ is given by $1/\lambda$; with $\lambda = 0.2$ the hypothetical mean is 5.0000

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

The mean of 1000 simulations of 40 exponential(0.2)s

```
mean(testResults$mean)
```

```
## [1] 4.979
```

This compares to the theoretical mean of an exponential distribution

```
(1/lambda)
```

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

Does the median = mean?

```
median(testResults$mean)
```

```
## [1] 4.974
```

2. Show how variable the data is and compare it to the theoretical variance of the distribution.

The variance of the 1000 means of 40 exponential(0.2)s

```
sd(testResults$mean)^2
```

```
## [1] 0.6323
```

This compares to the theorectical standard deviation of an exponential distribution
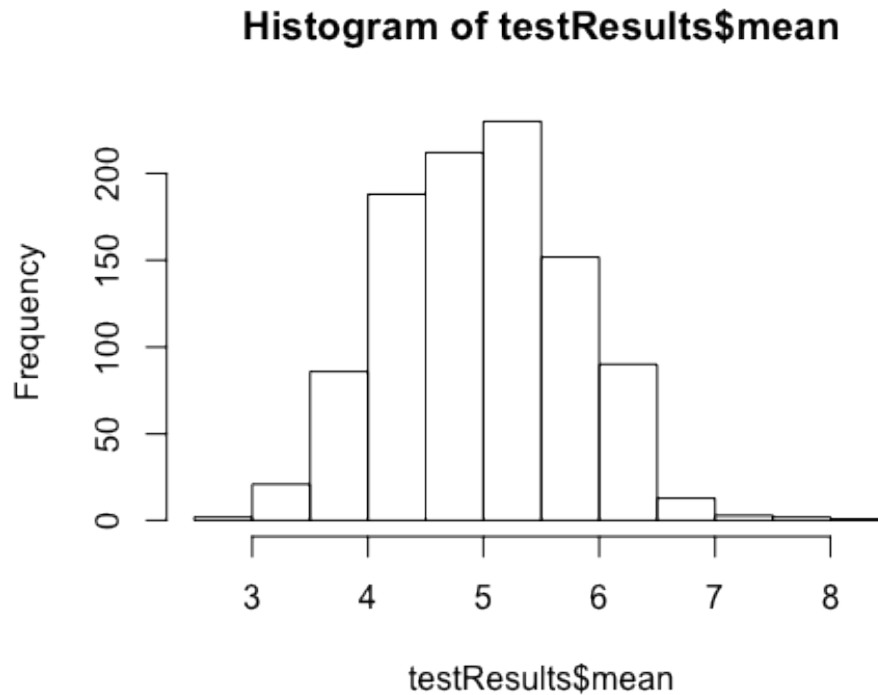
```
(1/lambda)^2/(n)
```

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

3. Show that the distribution is approximately normal.

Histogram of the mean of 40 exponential(0.2)s

```
hist(testResults$mean)
```

Histogram of testResults$mean

testResults$mean

4. Evaluate the coverage of the confidence interval

This is equal to the number of times the hypothetical mean is within each of the 1000 CIs divided by the number of simulations (to get a percentage)

```
length(testResults$coverageTest[testResults$coverageTest==1]) / simulations
```

[1] 0.922

We would expect the result to be close to 95%.