exponential_distribution

Kyle Scully

October 25, 2014

Introduction

The exponential distribution can be simulated in R with rexp(n, lambda) where lambda is the rate parameter. The mean of exponential distribution is 1/lambda and the standard deviation is also also 1/lambda. Set lambda = 0.2 for all of the simulations.

This project will simulate the distribution of acerages of 40 exponential(0.2)s.

The following properties will be discussed:

1. The center of the distribution
2. The variance of the distribution
3. The normality of the distribution
4. The coverage of the confidence interval for $\frac{1}{\lambda}$: $(\bar{X} \pm 1.96 \, \frac{S}{\sqrt{n}})$

Run Simulations

```
set.seed(123)
lambda = 0.2
n = 40
trials = 1000
simulation = matrix(rexp(n*trials, lambda), nrow = trials, ncol = n)
mean_of_each_trial = apply(simulation, 1, mean)
head(mean_of_each_trial)
```

```
## [1] 4.438 5.698 6.964 4.703 4.107 4.666
```

The center of the distribution

Our calculated center of the distribution from simulations is as follows:

```
mean(mean_of_each_trial)
```

```
## [1] 5.012
```

Our theoretical center of the distribution is:

$$\frac{1}{\lambda} = \frac{1}{0.2} = 5$$

Comparing our theoretical to our calculated we get a percent error of:

$$\frac{5.012-5}{5} = 0.0024 = 0.24\%$$

which means there is a high correlation between the centers of both the simulation and theoretical.

The variance of the distribution

To look into how much the distrubtion varies we will compare both calculated standard deviation and varience to the theoretical values, respectively.

Our calculated standard deviation of the distribution from simulations is as follows:

```
sd(mean_of_each_trial)
```

```
## [1] 0.7803
```

Our calculated variance of the distribution from simulations is as follows:

```
var(mean_of_each_trial)
```

```
## [1] 0.6088
```

Our theoretical standard deviation of the distribution is:

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

```
## [1] 0.7906
```

Our theoretical variance of the distribution is:

```
((1/lambda)/sqrt(n))^2
```

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

Comparing our theoretical to our calculated we get a percent error of:

Standard deviation: $\frac{.7906-.7803}{.7906} = 0.01302 = 1.30\%$

Variance: $\frac{.625-.6088}{.625} = 0.02592 = 2.59\%$

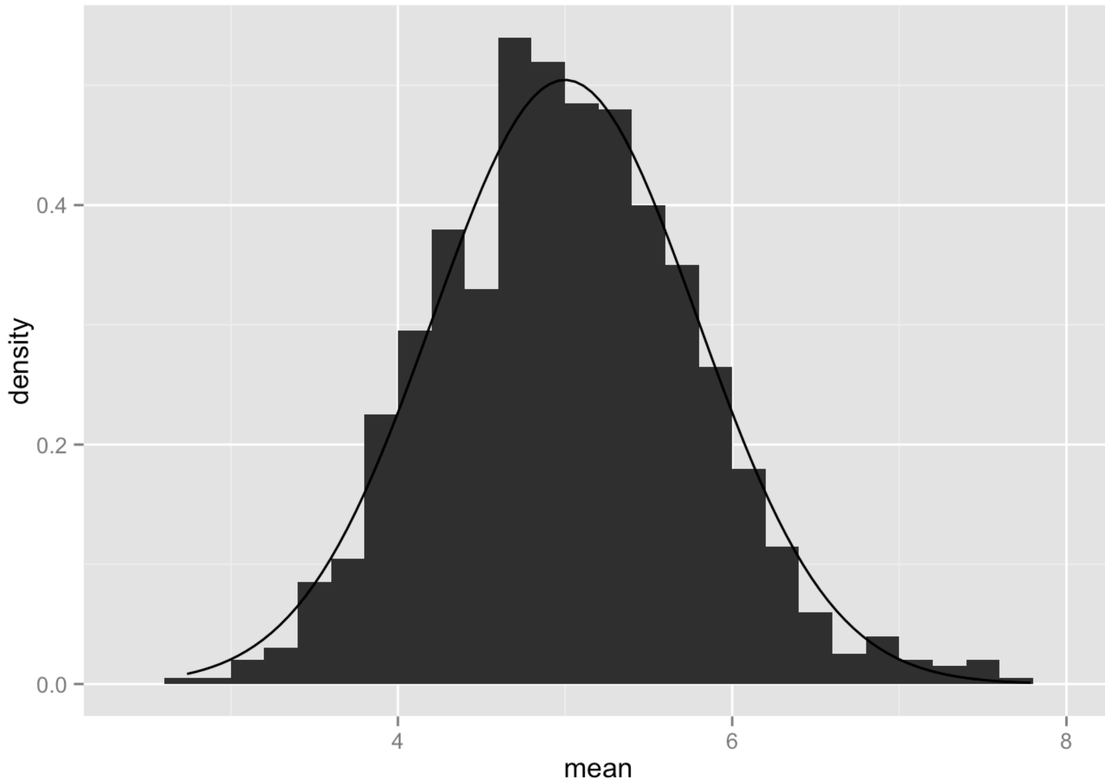
which means there is a high correlation between the each set of values.

The normality of the distribution

```
df = data.frame(mean = mean_of_each_trial, sd = apply(simulation, 1, sd)
)

library(ggplot2)

ggplot(data = df,
       aes(x=mean)) +
geom_histogram(aes(y=..density..),
               binwidth = lambda) +
stat_function(fun = dnorm,
              arg = list(mean = 5,
                         sd = .7906))
```



The figure above is a histogram of our simulation with the theoretical distribution curve overlayed. The simulation tends to be fairly normal.

The coverage of the confidence interval for $\frac{1}{\lambda}$: $(\bar{X} \pm 1.96 \frac{S}{\sqrt{n}})$

```
mean(mean_of_each_trial) + c(-1,1)*1.96*sd(mean_of_each_trial)/sqrt(tria
ls)
```

```
## [1] 4.964 5.060
```

The coverage of the 95% confidence interval for $\frac{1}{\lambda}$ is 4.964-5.060.