

Simulation Exercise

Amelia

2/11/2017

Overview

Simulations for 40 exponentials are ran for 1000 times and the mean and variance was compared to the theoretical values. The distribution is then compared to a normal distribution to show that it is approximately normal.

Data Processing

```
set.seed(100)
lambda = 0.2
n = 40

mns = NULL
for (i in 1 : 1000) mns = c(mns, mean(rexp(n, lambda)))
```

Show the sample mean and compare it to the theoretical mean of the distribution

```
samplemean = mean(mns)
realmean = 1/lambda
```

Sample mean is `samplemean` and theoretical mean is `realmean`.

Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution

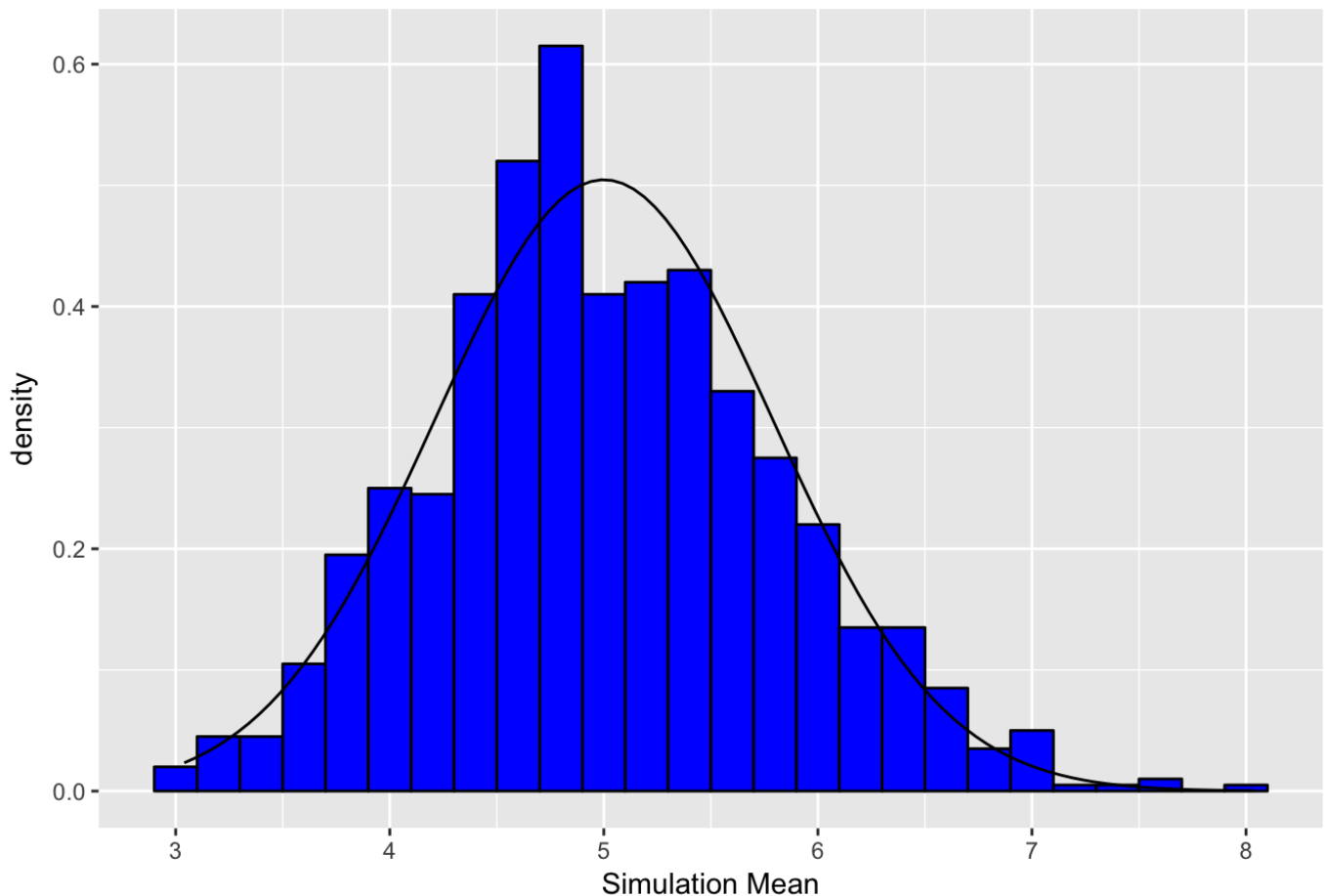
```
samplevar = var(mns)
realvar = (1/lambda)^2
```

Sample variance is NA and theoretical variance is NA.

Show that the distribution is approximately normal

```
library(ggplot2)
ggplot(data.frame(y=mns), aes(x=y)) +
  geom_histogram(aes(y=..density..), binwidth=0.2, fill="blue", colour = "black")
+
  stat_function(fun=dnorm, args=list(mean=lambda^-1, sd=(lambda*sqrt(n))^-1)) +
  labs(title="Plot of the Simulations", x="Simulation Mean")
```

Plot of the Simulations



Distribution is approximately normal.

Bonus: Difference between the distribution of a large collection of random exponentials and the distribution of a large collection of averages of 40 exponentials

```
lambda2 = 0.2
n2 = 1000

mns2 = NULL
for (i in 1 : 1000) mns2 = c(mns2, mean(rexp(n2, lambda2)))

library(tidyr)
library(dplyr)
total <- merge(mns, mns2)
total <- rename(total, s40 = x, s1000 = y)
total2 <- gather(total, simulations)

g<- ggplot(data=total2, aes(x=value, fill=simulations)) + geom_density(alpha=.3) +
  labs(title="Plot of the Simulations", x="Simulation Mean")
print(g)
```

Plot of the Simulations

