Statistical Inference: Simultation

Seth Dimick

12/24/2017

Here we will explore the theoretical properties of an exponential distribtion versus a simulation. Given an exponential distribution with a rate of 0.2, we'll explore if the mean is equal to the 1/rate (5), if the standard deviation is also equal to 1/rate (5), and if the distribution of averages is normally distributed.

Let's record the mean of standard deviation of one thousand simulated exponentials with a rate of 0.2 and 40 observations and analyze the results.

```
#Generate Simulated Data
sim <- data.frame(mean = NULL, sd = NULL)
for (i in 1 : 1000) {
  obs <- rexp(40, rate = 0.2)
  stat <- data.frame(mean = mean(obs), sd = sd(obs))
  sim <- rbind(sim, stat)
}
```

1.1 Exponential Mean

Given that our simulated data was created with a rate of 0.2, we should see a mean of 1/0.2, or 5. We observe a mean of our simulated means to be 4.99, which supports the theoretical mean with some error.

1.2 Exponential Variance (Standard Deviation)

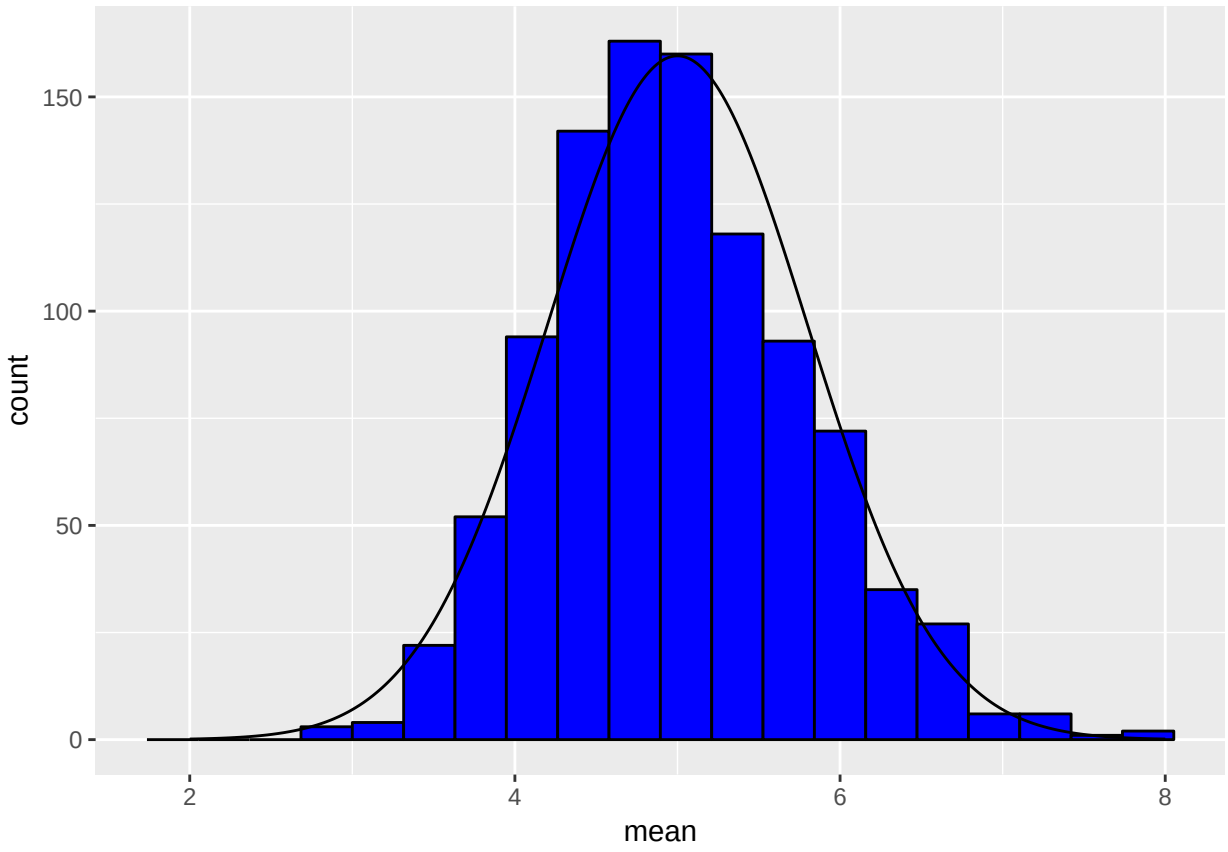
Given that our simulated data was created with a rate of 0.2, we should see a standard deviation of 1/0.2, or 5. We observe a mean of our simulated standard deviations to be 4.87, which supports the theoretical standard deviation with some error.

1.3 Average of Exponentials

We can finally overlay a normal distribution over the distribution of our simulated means. Doing so we observe that the distribution of the means is indeed, approximately normal.

```
x <- seq(2,8,length=1000)
y <- dnorm(x,mean=5, sd=0.8)
normal <- data.frame(x=x, y=y*320)

ggplot(sim, aes(x=mean)) +
  geom_histogram(bins = 20, fill="blue", color="black") +
  geom_line(data = normal, aes(x=x, y=y))
```


150
100
50
0
count
2
4
6
8
mean