Simulation Exercise

Statistical Inference Class Project part 1

Ben Garthwaite

27 September 2015

Synopsis

An investigation of the exponential distribution compared to the Central Limit Theorem.

Simulations

For 1000 simulations, what is the mean of 40 exponentials with lambda equal to 0.2:

```
set.seed(3)
nosim <- 1000
n <- 40
lambda <- 0.2
simulations <- replicate(nosim,rexp(n,lambda))
simulations_mean <- mean(simulations)
dist <- apply(simulations, 2, mean)
```

Sample Mean versus Theoretical Mean

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

```
# Return Sample Mean
simulations_mean
```

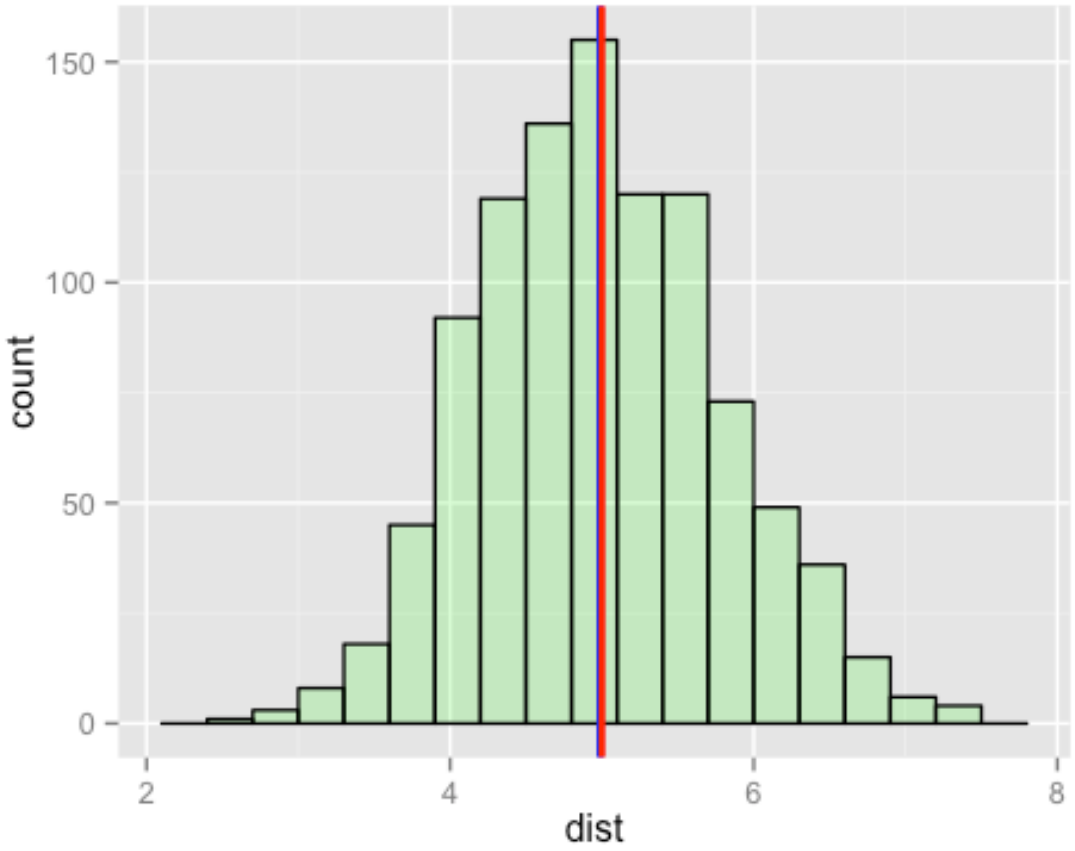
```
## [1] 4.98662
```

```
# Calculate Theoretical Mean (1/lamda)
theo_mean <- 1/lambda
theo_mean
```

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

```
require(ggplot2)
dat <- data.frame(dist)
g <- ggplot(dat, aes(x = dist))
g <- g + geom_histogram(alpha=.2, binwidth=.3, colour="black", fill="gr
een")
g <- g + geom_vline(xintercept = simulations_mean, colour="blue", size=
1)
```

```
g <- g + geom_vline(xintercept = theo_mean, colour="red", size=1)
g
```



In the chart above, the sample mean is represented by a blue bar and the theoretical mean is represented by a red bar, which demonstrates how close the sample mean of 4.9866197 is to the theoretical mean 5.

Sample Variance versus Theoretical Variance

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

```
# Calculate Sample Variance
simulations_sd <- sd(dist)
simulations_sd

## [1] 0.7947823

simulations_sd^2

## [1] 0.6316789
```

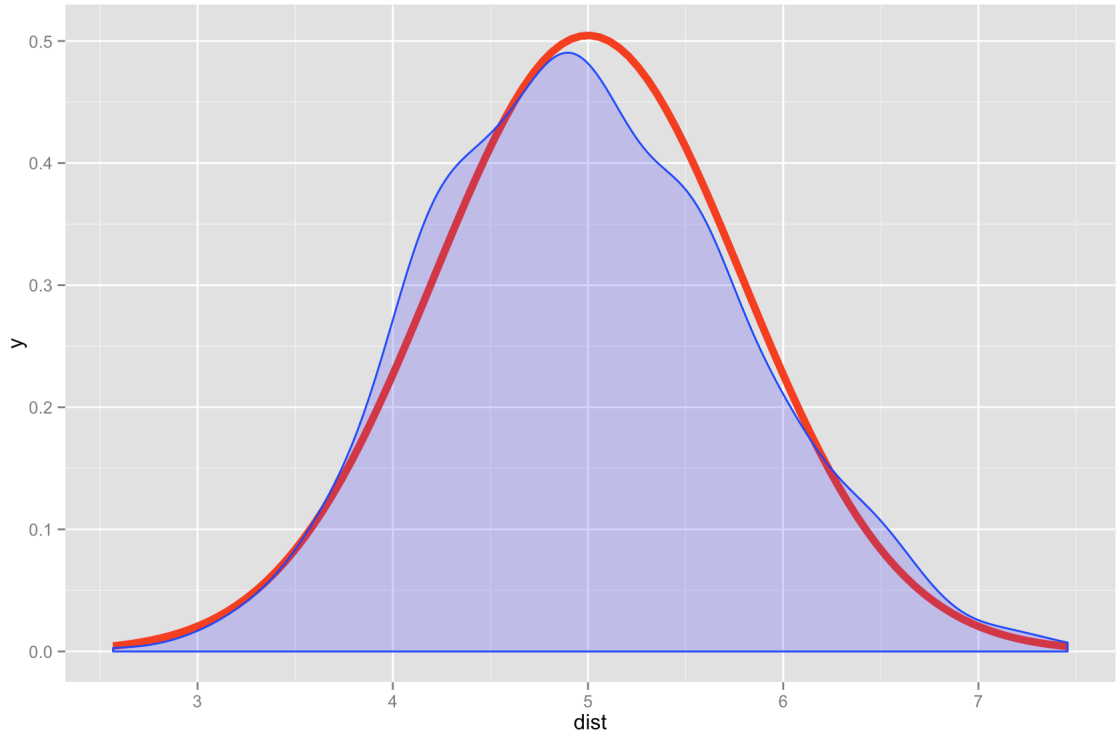
```
# Calculate Theoretical Variance
theo_sd <- (1/lambda)/sqrt(n)
theo_sd

## [1] 0.7905694

theo_sd^2

## [1] 0.625
```

To confirm the calculations above the sample distribution standard deviation is 0.7947823 and the variance is 0.6316789. Again this is close to the theoretical values: standard deviation of 0.7905694 and variance of 0.625.

Distribution

Show that the distribution is approximately normal.

```
g <- ggplot(dat, aes(x=dist))
g <- g + stat_function(fun=dnorm,args=list(mean=theo_mean,sd=theo_sd),c
olor="red", size=2)
g <- g + geom_density(alpha=.2, fill="blue", colour="blue")
g
```



The chart above plots the sample distribution in blue, with a thick red line showing the normal distribution. It demonstrates that the sample distribution is approximately normal.