Simulation of a Convergence to the Normal Distribution

Jim Callahan

September 17, 2015

Overview

In this project we demonstrate how averages of an asymetric distribution (the possion distribution) converge to a symetric normal distribution.

This convergence is an application of the Central Limit Theorem (CLT). "For our purposes, the CLT states that the distribution of averages of iid variables, properly normalized, becomes that of a standard normal as the sample size increases" slide from Dr. Brian Caffo's *"Statistical Inference"* class slide 8/20 "A Trip to Asymptopia" in Asymptopia.pdf

Simulations

Include English explanations of the simulations you ran, with the accompanying R code. Your explanations should make clear what the R code accomplishes.

In this project you will investigate the exponential distribution in R and compare it with the Central Limit Theorem. 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 1/lambda. **Set lambda = 0.2** for all of the simulations.

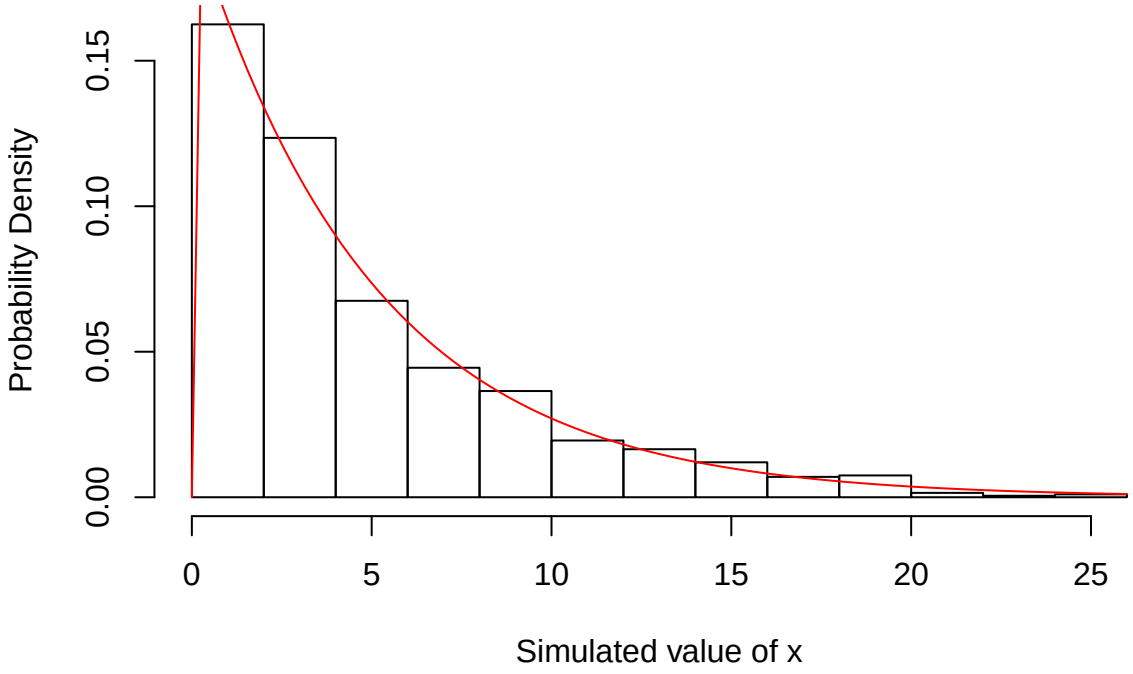
You will investigate the distribution of averages of 40 exponentials. Note that you will need to do a thousand simulations.

Get code from ExponentialSim.R

```
# Simulate 1000 random exponentials
x <- rexp(1000, rate = 0.2)

# Plot the empirical results of the simulation
# Use density instead of frequency so the theoretical curve will fit.
hist(x, freq=F,
     main = "Histogram of 1,000 random exponentials
             \n with overlay of theoretical density",
     xlab = "Simulated value of x",
     ylab = "Probability Density"
     )


# Overlay a plot a theoretical exponential density curve
curve(dexp(x, rate = 0.2), add=T, col="Red")
```



Sample Mean versus Theoretical Mean

Include figures with titles. In the figures, highlight the means you are comparing. Include text that explains the figures and what is shown on them, and provides appropriate numbers.

Sample Variance versus Theoretical Variance

Include figures (output from R) with titles. Highlight the variances you are comparing. Include text that explains your understanding of the differences of the variances.

Distribution

Via figures and text, explain how one can tell the distribution is approximately normal.

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

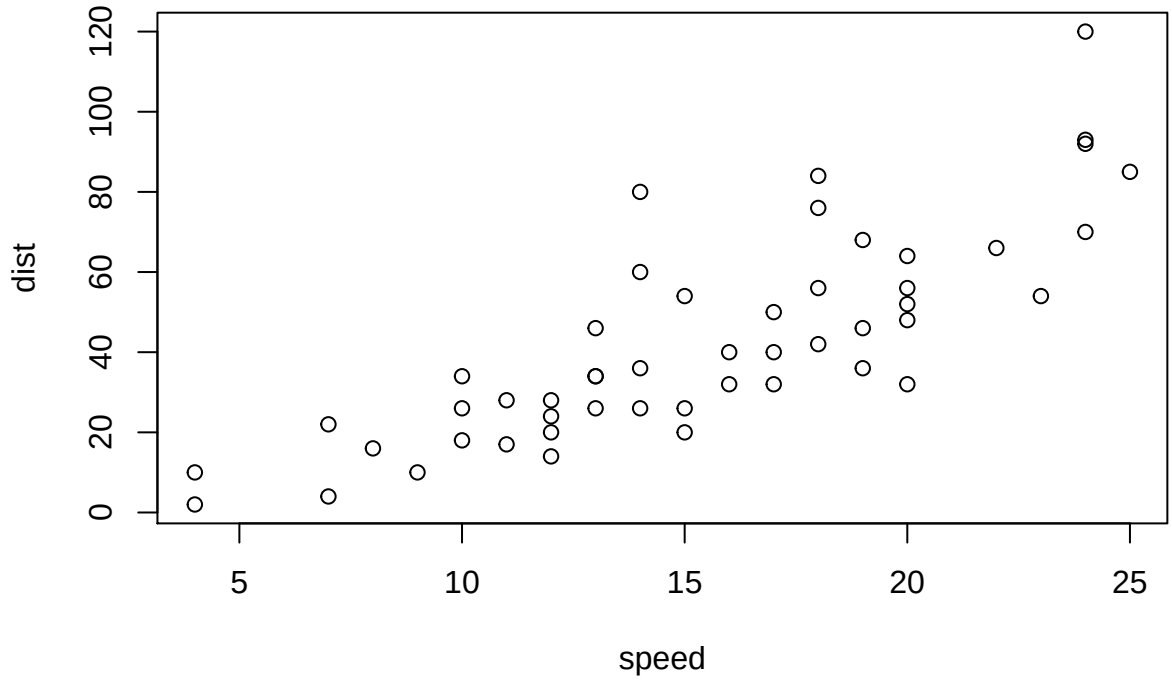
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed           dist
##  Min.   : 4.0   Min.   :  2.00
```

```
##  1st Qu.:12.0   1st Qu.: 26.00
##  Median :15.0   Median : 36.00
##  Mean   :15.4   Mean   : 42.98
##  3rd Qu.:19.0   3rd Qu.: 56.00
##  Max.   :25.0   Max.   :120.00
```

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.