Investigate Various Exponential Distribution Statistics

YeongWei

Saturday, November 21, 2015

Overview

This document aims to investigate the various statistical properties for Exponential Distribution by comparing the sample and theoretical statistical properties.

Simulations

The R Code below attempts to produce Exponential Distributions of 40, 80 number of observations for 1000 number of simulations with λ as 0.2 then for each simulation compute the mean and variance.

```
run <- function(lambda, numberOfSimulations, numberOfObservations) {
  statisticsByGroup <- data.frame(Group = c(), Mean = c(), Variance = c());
  for (x in numberOfObservations) {
    meanAccumulator <- c();
    varianceAccumulator <- c();
    for (n in 1 : numberOfSimulations) {
      expSampleDistribution <- rexp(x, lambda);
      meanAccumulator <- c(meanAccumulator, mean(expSampleDistribution));
      varianceAccumulator <- c(varianceAccumulator, var(expSampleDistribution));
    }
    group <- rep(x, numberOfSimulations);
    groupMeanVariance <- cbind(Group = group, Mean = meanAccumulator,
                               Variance = varianceAccumulator);
    statisticsByGroup <- rbind(statisticsByGroup, groupMeanVariance);
  }
  statisticsByGroup$Group <- as.factor(statisticsByGroup$Group);
  statisticsByGroup;
}

set.seed(10000);
lambda <- 0.2; numberOfSimulations <- 1000; numberOfObservations <- c(40, 80);

statisticsByGroup <- run(lambda, numberOfSimulations, numberOfObservations);
```

Theoretical Statistics

Given that, mean is $\mu = 1 \ / \ \lambda$, standard deviation is $\sigma = 1 \ / \ \lambda$, variance is σ^2 and $\lambda = 0.2$. The following attempts to compute the theoretical mean, standard deviation and variance.

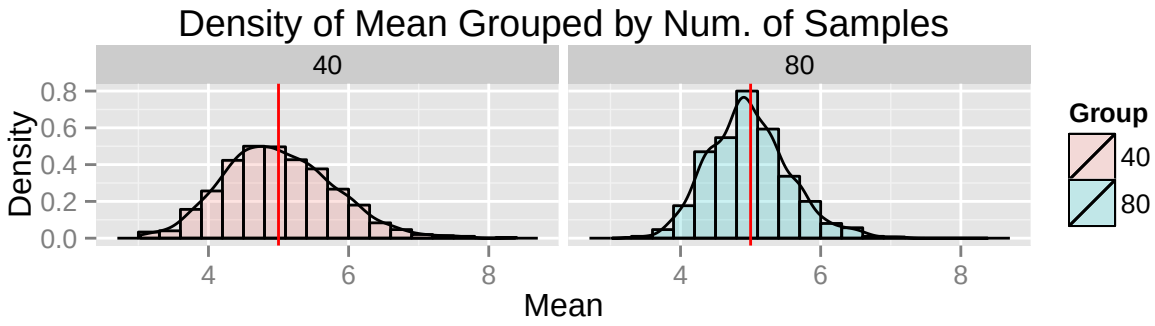
Mean, Standard Deviation and Variance

```
## [1]  5  5 25
```

Statistical Inferences

This section attempts to investigate the sample mean and variance against the theoretical ones.

Sample Mean Versus Theoretical Mean



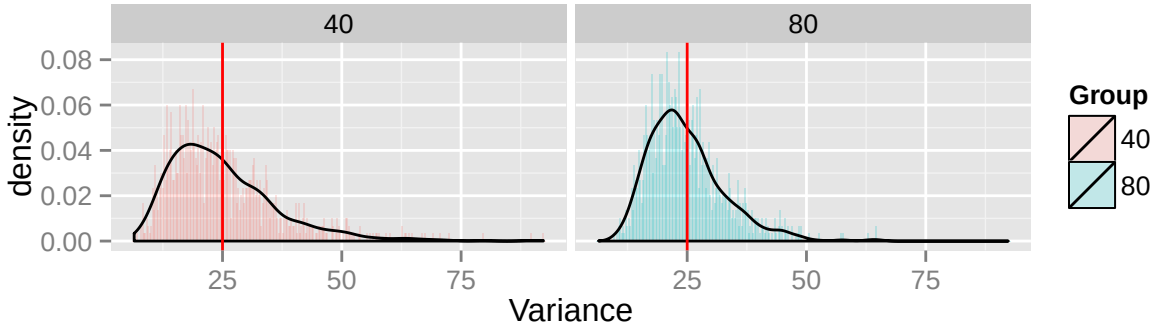
1. The average sample mean for the group of 40 and 80 number of observation,

```
c(mean(statisticsByGroup[statisticsByGroup$Group == 40, "Mean"]),
  mean(statisticsByGroup[statisticsByGroup$Group == 80, "Mean"]));
```

```
## [1] 5.005990 4.978395
```

2. With more samples, the sample mean values are more likely to converges to the theoretical mean value, 5 represented by the red vertical line.

Sample Variance Versus Theoretical Variance



1. The average sample variance for the group of 40 and 80 number of observation,

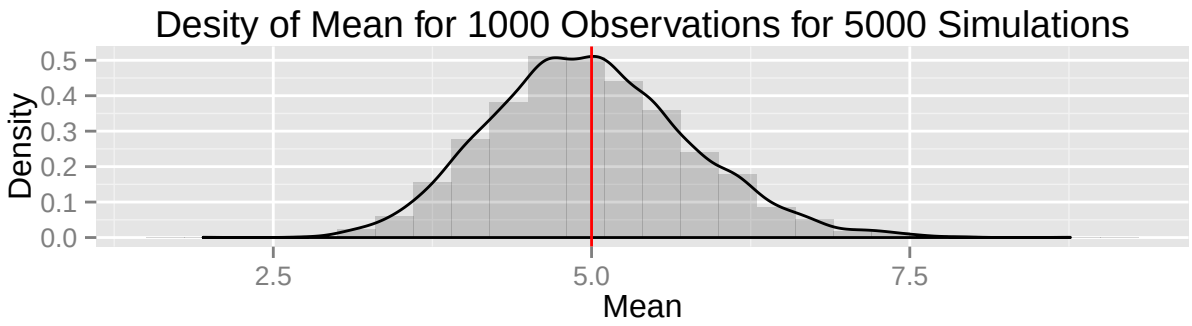
```
c(mean(statisticsByGroup[statisticsByGroup$Group == 40, "Variance"]),
  mean(statisticsByGroup[statisticsByGroup$Group == 80, "Variance"]));
```

```
## [1] 25.05950 24.76162
```

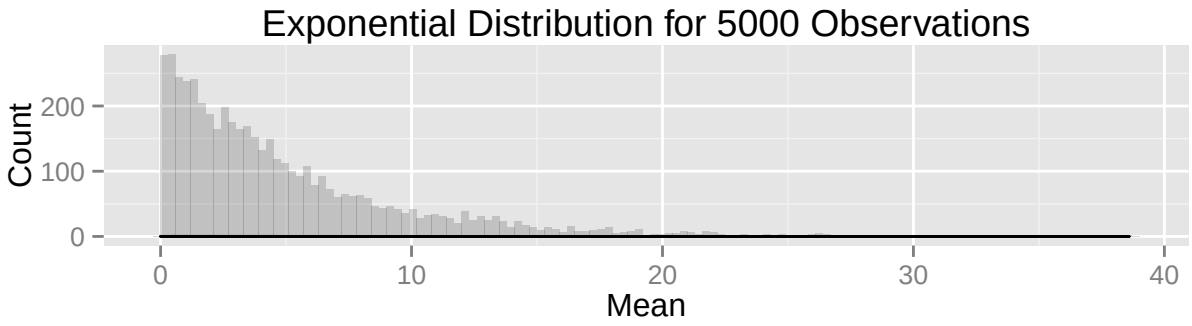
2. With more samples, the sample variance values are more likely to converges to the theoretical variance value, 25 represented by the red vertical line.

Further Analysis

Sample Mean for 40 Observations for 5000 simulations



Exponential Distribution of 5000 Observations with $\lambda = 0.2$



Summary

1. As the sample size increases, the mean of samples converges to the theoretical mean and the sample variance converges to theoretical variance as described by the Law of Large Numbers.

2. Exponential Distribution is clearly not normally distributed. However according to Central Limit Theorm, as the sample size increases, the averages of samples converges to normal distribution.