Proving the Central Limit Theorem

Proving the Central Limit Theorem

dchen71

Overview

This report will compare the expontential distrbution against the normalized central limit theorem. The report will show how to tell if a distribution is approximately normal.

Simulations

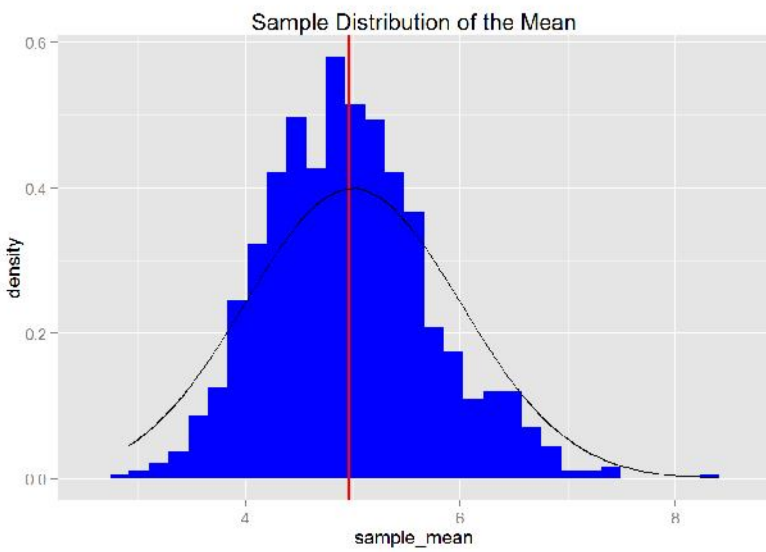
The following code is the code to create the simulation data. The code will setup an expontential distribution as well as its assocaited means and variability. The simulations are running expotential distrubtions of 40 times with a rate of 0.2 The means and variability are stored for future usage. It will also save the overall mean and variance of the simluations done.

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.1.3
```

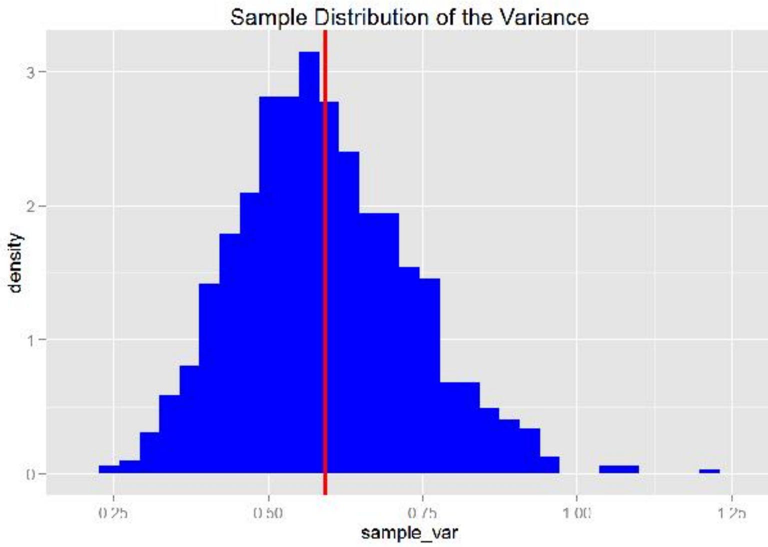
```
set.seed(10)

lambda = .2

sample_mean = 0
sample_var = 0
for(i in 1:1000){
    sample_mean[i] = mean(rexp(40,lambda))
}
for(i in 1:1000)
    sample_var[i] = var(sample(sample_mean, 40, replace = TRUE))


sample.mean = mean(sample_mean)
sample.var = var(sample_mean)

sample_mean = as.data.frame(sample_mean)
sample_var = as.data.frame(sample_var)
```

Sample Means vs Theoretial Mean



The above figure shows the distribution of sampled means against the population mean. As can be seen in the plot above, the distribution of the sample mean is close to the actual theoretical value of the population mean of 'r sample_mean'. The black line is the overlaid normal distribution. The exponential distribution roughly follows the normal distribution.

Sample Variance vs Theoretical Variance



The above figure shows the distribution of sampled variances against the theoretical variance of 'r sample.var'. The theoretical variance is ((1/lambda)^2)/n which is 0.625. The variance of the samples is 'r sample.var' which is close to the theoretical value.