Course Project

Overview

This report is broken into two parts, each doing separate analyses. The first investigates the exponential distribution in R and compare it with the Central Limit Theorem. Part two is an analysis of the ToothGrowth data in the R datasets package. ## Part 1 # Simulations The simulation run was of the exponential distribution. First the variables used in the simulation were each defined:

```
set.seed(21)
n <- 40
lambda <- 0.2
sim <- 1000
```

The exponential distribution is then simulated 1000 times by the rexp function, where lambda is the rate parameter.

```
dist<-matrix(rexp(n*sim, lambda), sim)
```

Finally, the average of exponetial is calculated:

```
rowMeans <- apply(dist,1,mean)
```

Sample Mean Vs. Theoretical Mean

The mean of the sample was calaculated to be 4.98144:

```
sampleMean <- mean(rowMeans)
sampleMean
```

```
## [1] 4.98144
```

This sample mean is lower than the theoretical mean of 5, calculated by dividing 1/lambda:

```
popMean<-1/lambda
popMean
```

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

Sample Variance vs. Theoretical Variance

The variance of the sample was calculated to be .604645

```
sampleVar <- var(rowMeans)
sampleVar
```

```
## [1] 0.604645
```

This was slightly lower than the theoretical variance, calculated to be .625:

```
popSd<-((1/lambda) * (1/sqrt(n)))
popVar<- popSd^2
popVar
```

```
## [1] 0.625
```

Distribution

When vieiwing a histogram of the density distribution of the simulation, one can see the it follows an approximately normal distribution. The exception to this is that the higher values occur just outside the median,and the sample has an extended right tail.

```
data<-data.frame(rowMeans)
xfit <- seq(min(rowMeans), max(rowMeans), length=100)
yfit <- dnorm(xfit, mean=1/lambda, sd=(1/lambda/sqrt(n)))
hist(data$rowMeans,breaks=n,prob=T,col="blue",xlab = "means",main="Density Distribution of Average
  of Exponentials",ylab="density")
lines(xfit, yfit, pch=22, col="black", lty=5)
```


Density Distribution of Average of Exponentials