Statistical Inference Project

Mayank Pundhir

October 26, 2014

Part 1 of the Project: Simulation Exercise

In this simulation, the distribution of averages of 40 exponentials with lambda 0.2 has been investigated. No of Simulations are 4000. All questions were answered.

1. Show where the distribution is centered at and compare it to the theoretical center of the distribution.

Ans: The distribution generated in simulation was centered at 4.9904 whereas theoretical mean of the distribution should be at 5. The difference was very small and non significant.

```
t.test(Y)$conf.int
```

```
## [1] 4.966 5.015
## attr(,"conf.level")
## [1] 0.95
```

The Confidence Interval of mean of observation contains the theoretical mean.

2. Show how variable it is and compare it to the theoretical variance of the distribution.

Ans: The distribution generated in simulation was centered at 0.8004 whereas theoretical mean of the distribution should be at 0.7906. This difference was also non significant.

```
sqrt((sim-1)*S^2/qchisq(c(0.975,0.025),sim-1))
```

```
## [1] 0.7833 0.8184
```

The Confidence Interval of standard deviation of observation contains the theoretical deviation.

3. Show that the distribution is approximately normal.

Ans: QQ normal plot of the variable was plotted which show normality of the data. Shapiro-Wilk normality test was also done on the data and the Shapiro-Wilk statistic was and p-value came out to be which confirmed the result from the plot.

4. Evaluate the coverage of the confidence interval for $1/\lambda : \bar{X} \pm 1.96\frac{S}{\sqrt{n}}$.

Ans: Confidence Interval(95%) of the variable was 3.4505 to 6.5495. Coverage of this confidence interval was calculated which came out to be 0.9448.

Appendix

Part 1:

Generation of Data

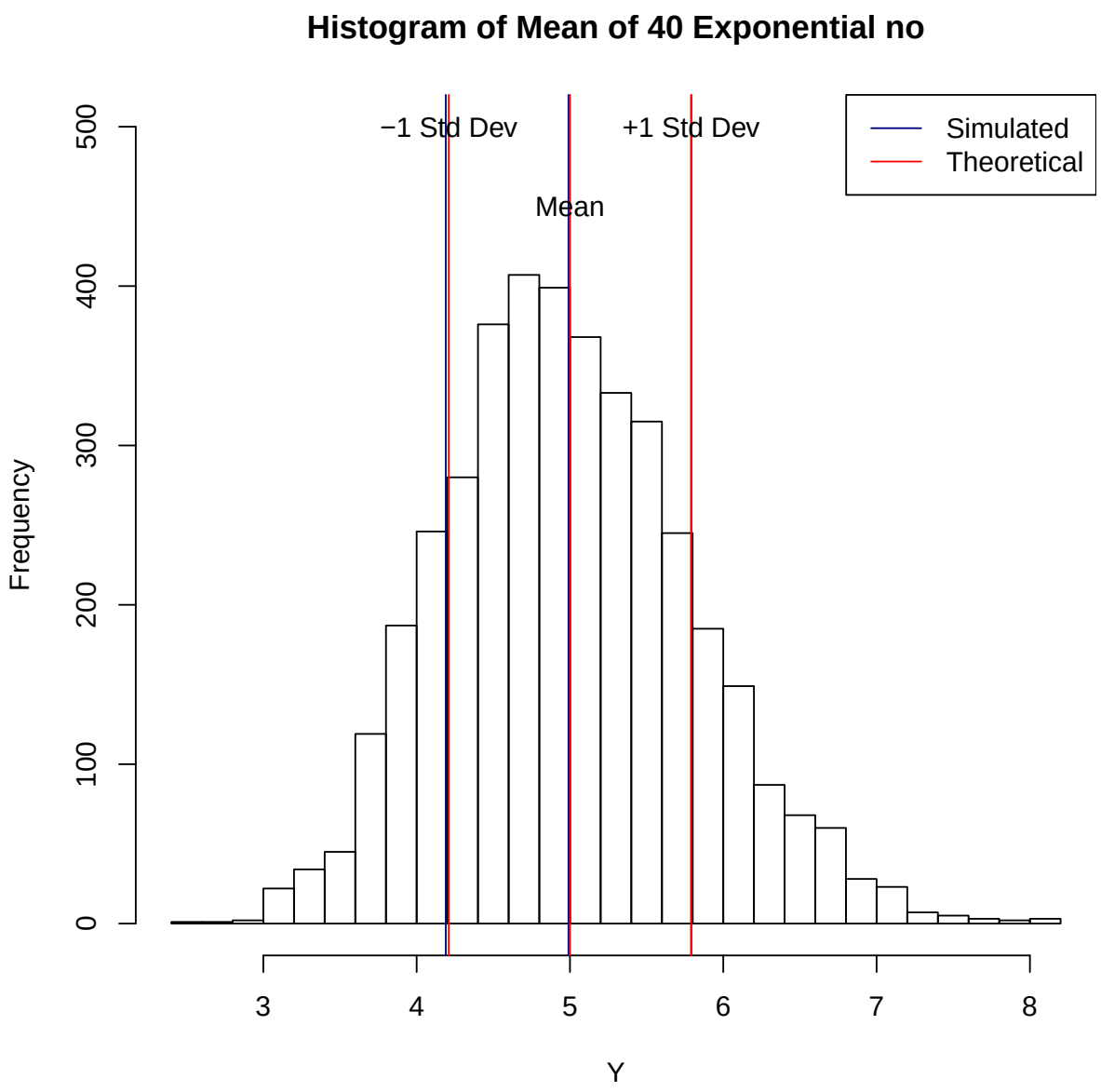
```
set.seed(100);lambda<-0.2;sim<-4000;obs<-40;
X=data.frame(matrix(NA, nrow=sim, ncol=obs)) ##Empty dataframe
for(i in 1:sim){  X[i,]<-rexp(obs,lambda)}##Getting random no.
Y<-rowMeans(X) ##Mean of 40 exponentials
```

Observed and Theoretical mean and standard variance

```
mu<-mean(Y);S<-sd(Y) ##Observed Mean and Standard deviation of Y
mean<-1/lambda;sd<-(1/lambda)/sqrt(obs) ##Theoretical mean and std. of Y
```

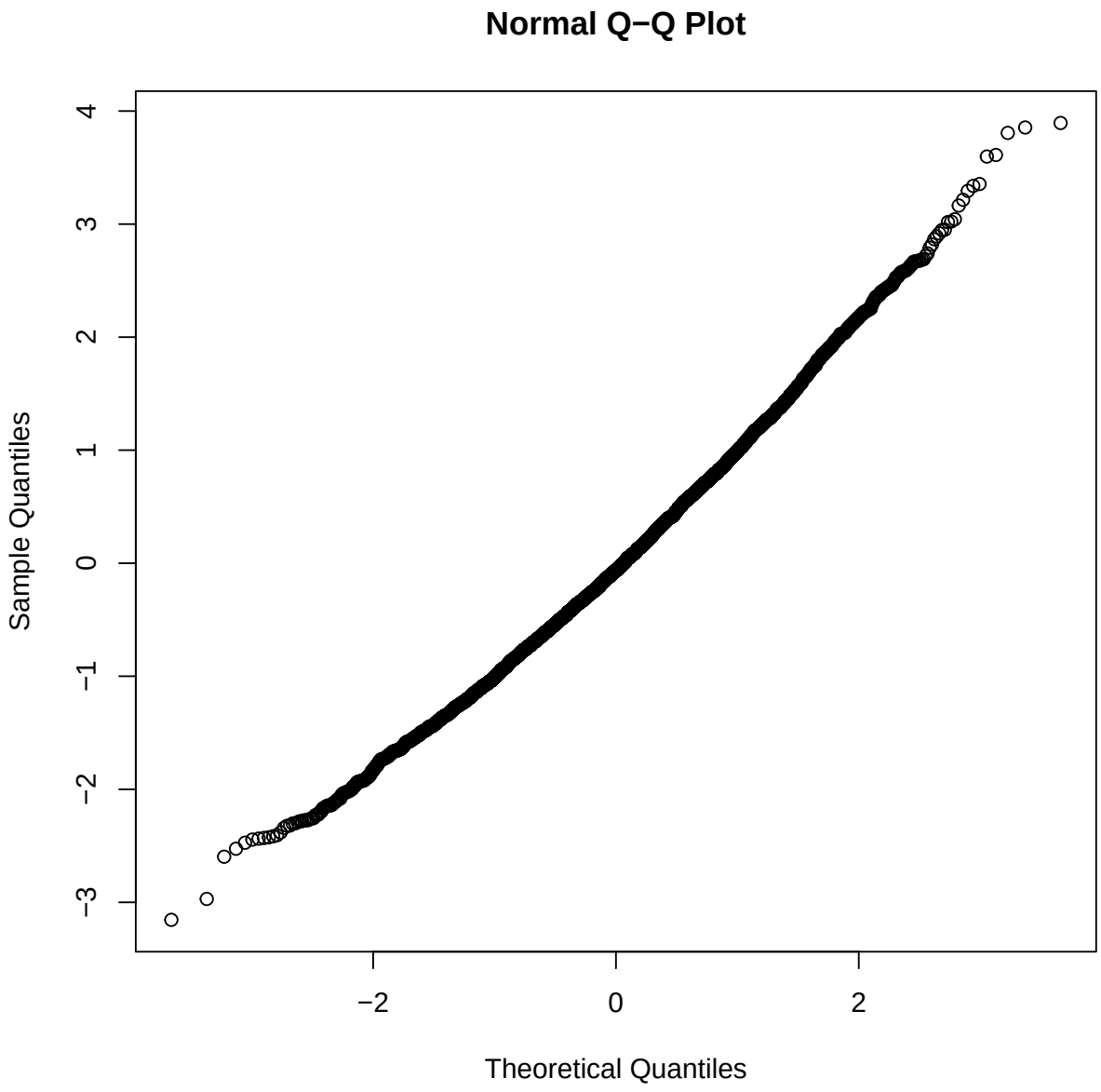
Histogram

```
hist(Y,"Scott",main="Histogram of Mean of 40 Exponential no",ylim=c(0,500))
abline(v=mu,col="navy");abline(v=c(mu-S,mu+S),col="navy")
abline(v=mean,col="red");abline(v=c(mean-sd,mean+sd),col="red")
text(mean-sd,500,"-1 Std Dev");text(mean+sd,500,"+1 Std Dev");text(mean,450,"Mean")
legend("topright",c("Simulated","Theoretical"),lty=1,col=c("navy","red"))
```



Normality

```
qqnorm((Y-mu)/S)
```



Confidence Intervals and Coverage

```
conf_int<-mean+c(-1,1)*1.96*sd
coverage<-(conf_int[1]<Y & Y<conf_int[2]);mean(coverage)
```