Statistical Inference Course Project

Lee 25 Jan 2015

Overview: In a few (2-3) sentences explain what is going to be reported on.

Part 1: Simulations:

```
set.seed(17) # set seed
this_lambda <- 0.2 # set lambda
this_sample_size <- 40 # set sample size
# simulate with 40 exponential samples 1000 times
sim.set <- replicate(1000, rexp(this_sample_size, this_lambda))
# get the mean of each simulation
sim.mean <- apply(sim.set, 2, mean)
```

Sample Mean versus Theoretical Mean:

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

```
# mean of the means of the simulations and the theoretical center
c(mean(sim.mean), 1/this_lambda)
```

```
## [1] 4.962 5.000
```

The center of distribution of the mean of 40 exponentials is close to the theoretical center of the distribution. This is supported with the graph below.

```
hist(sim.mean, probability=T, breaks=30, main="Distribution of the mean of 40 exponentials in 1,000 simu
abline(v=mean(sim.mean), col="red")
abline(v=1/this_lambda, col="blue", lty=2)
lines(density(sim.mean), col="red")
x<-seq(min(sim.mean), max(sim.mean),1000)
curve(dnorm(x, mean=mean(sim.mean), sd=sd(sim.mean)), add=TRUE, col="blue", lty=2)
legend('topright', legend=c("Simulation", "Theoritcal"), lty=c(1,2), col=c('red', 'blue'))
```

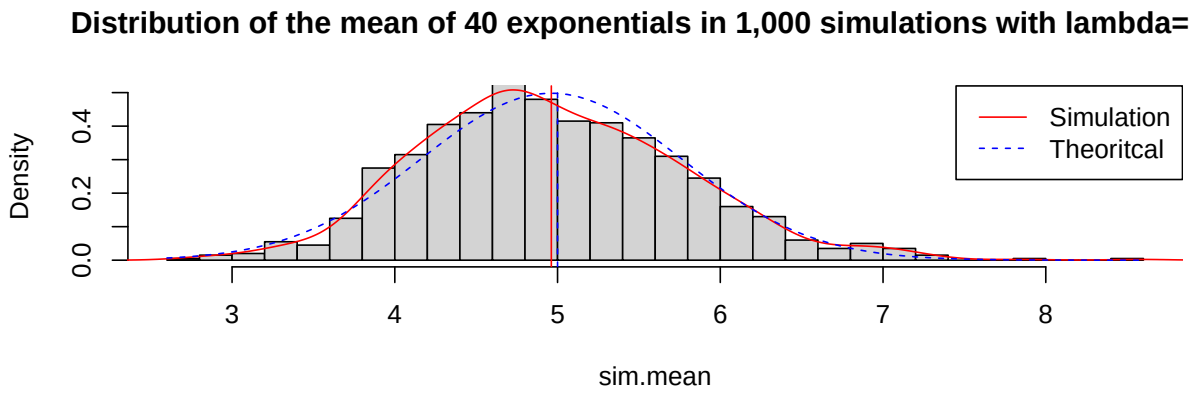

Figure 1

Sample Variance versus Theoretical Variance:

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

```
# Compare the standard deviation and variance in the sample and theory
matrix(c(sd(sim.mean), (1/this_lambda)/sqrt(this_sample_size),var(sim.mean),
         1/((this_lambda^2) * this_sample_size)),ncol=2,byrow=TRUE,
       dimnames=list(c("Standard deviation","Variance"), c("Sample", "Theory")))
```

```
##                    Sample Theory
## Standard deviation 0.8005 0.7906
## Variance           0.6408 0.6250
```

The standard deviation and the variance in the sample and theory are close. The standard devidation in the samples are slightly greater than that of the theory as shown in Figure 2.

```
sample_sd <- (1/this_lambda)/sqrt(this_sample_size)
this_sd <- sd(sim.mean)
hist(sim.mean, probability=T, breaks=30, main="Distribution of the mean of 40 exponentials in 1,000 simu
x<-seq(min(sim.mean), max(sim.mean),1000)
curve(dnorm(x, mean=mean(sim.mean), sd=sd(sim.mean)), add=TRUE, col="blue", lty=2)
legend('topright', legend=c("Simulation", "Theoritcal"), lty=c(1,2), col=c('red', 'blue'))
lines(density(sim.mean), col="red")
abline(v=mean(sim.mean), col="red")
abline(v=mean(sim.mean)+sample_sd, col="red")
abline(v=mean(sim.mean)-sample_sd, col="red")
abline(v=mean(sim.mean)+(2*sample_sd), col="red")
abline(v=mean(sim.mean)-(2*sample_sd), col="red")
abline(v=mean(sim.mean)+(3*sample_sd), col="red")
abline(v=mean(sim.mean)-(3*sample_sd), col="red")
abline(v=1/this_lambda, col="blue", lty=2)
abline(v=(1/this_lambda)+this_sd, col="blue", lty=2)
abline(v=(1/this_lambda)-this_sd, col="blue", lty=2)
abline(v=(1/this_lambda)+(2*this_sd), col="blue", lty=2)
abline(v=(1/this_lambda)-(2*this_sd), col="blue", lty=2)
```

```
abline(v=(1/this_lambda)+(3*this_sd), col="blue", lty=2)
abline(v=(1/this_lambda)-(3*this_sd), col="blue", lty=2)
```

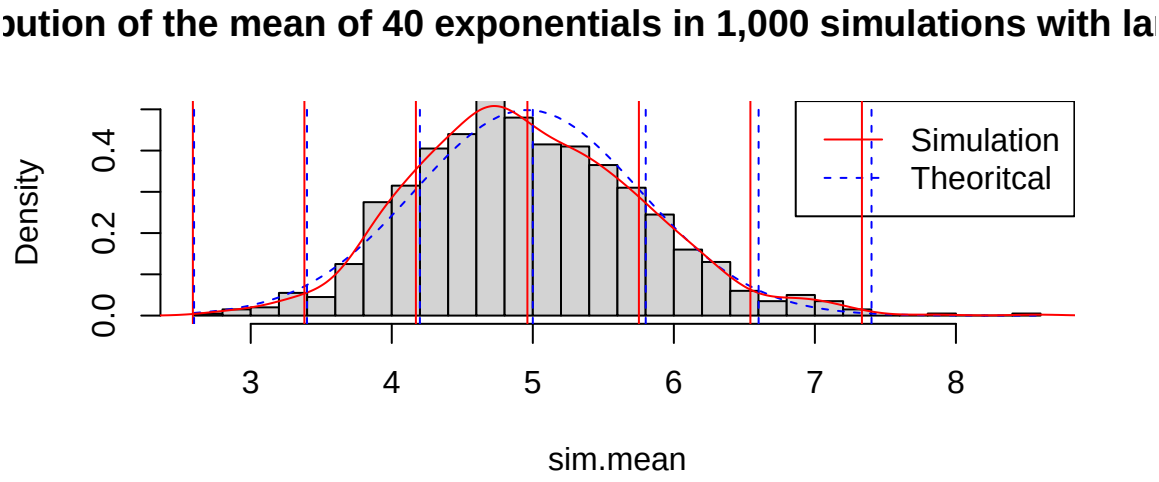

Figure 2

Distribution:

Central limit theorem suggests the sample means follow normal distribution. Figure 1 shows the simulated sample means are close to a normal distribution. Figure 3 compares the meas of 40 exponentials of the 100 simulations to a normal distribution and suggests the distribution of the means are very close to a normal distribution.

```
# Compare the distribution of averages of 40 exponentials to a normal distribution
qqnorm(sim.mean, main="Distribution of the means of 40 exponentials")
qqline(sim.mean, col = "red")
legend('bottomright', legend="Normal Distribution", lty=1, col="red")
```

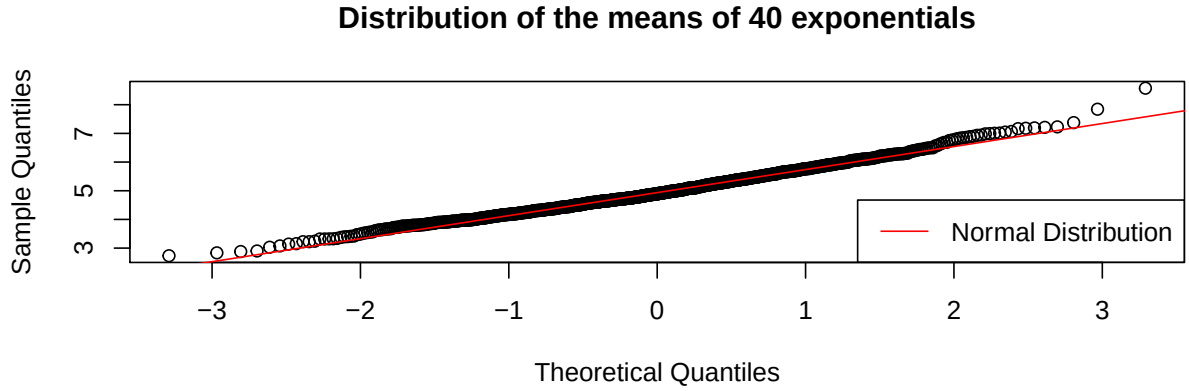

Figure 3