Statistical Inference Project Part 1

James Hancock

July 30, 2016

Exercise

The exponential distribution can be simulated in R with rexp(n, lambda) where lambda is the rate parameter. The mean of exponential distribution is 1/lambda and the standard deviation is also 1/lambda. Set lambda = 0.2 for all of the simulations. You will investigate the distribution of averages of 40 exponentials. Note that you will need to do a thousand simulations.

Illustrate via simulation and associated explanatory text the properties of the distribution of the mean of 40 exponentials. You should 1. Show the sample mean and compare it to the theoretical mean of the distribution. 2. Show how variable the sample is (via variance) and compare it to the theoretical variance of the distribution. 3. Show that the distribution is approximately normal.

Question 1

set seed

```
set.seed(2)
}
```

set Lamda

```
lamda <- 0.2
}
```

simulations

```
sim <- 1000
}
```

samples

```
n <- 40
}
```

Simulate & Take Mean

```
simexp <- replicate(sim, rexp(n,lamda))
meanexp <- colMeans(simexp)
}
```

I now need to compare the simulated mean with the theoretical mean.

simulated mean

```
smean <- mean(meanexp)
print(smean)
}
```

5.016356

theoretical mean

```
tmean <- 1/lamda
print(tmean)
}
```

5

Graph to compare

```
hist(meanexp, breaks = 20, xlab = "Mean", main = "Histogram of 1000 Simulated Exponential Means", col =
abline(v = smean, col = "red")
}
```

Question 1 The simulated mean is 5.01, compared to the calculated theoretical mean of 5.

Question 2

Simulated Standard Deviation and Variance

```
sstdev <- sd(meanexp)
svar <- sstdev^2
print(svar)
}
```

0.6691305

Theoretical Standard Deviation and Variance

```
tstdev <- (1/lamda)/sqrt(n)
tvar <- tstdev^2
print(tvar)
}
```

0.625

Question 2 answer: The simulated variance is 0.6691 compared to the theoretical variance of 0.625

Question 3

Plot Histogram

```
hist(meanexp, breaks = 40, xlab = "Mean", main = "Comparison to a Normal Distribution", col = "beige")
}
```

Add the Theoretical Normal Distribution Line

```
xfit <- seq(min(meanexp), max(meanexp), length = 100)
yfit <- dnorm(xfit, mean = 1/lamda, sd = 1/lamda/sqrt(n))
lines(xfit, yfit*100, lty=2)
}
```

As the graph shows, the distribution is closely matched with a normal distribution