Statistical Inference: Peer Assessment 1

We begin by examing the exponential distribution. Figure 1 shows the histogram generated from looking at the distribution of 1000 averages of 40 exponentials. The center appears to be slightly less than 5 which is in close agreement with the theoretical center of 5 ($1/0.2 = 5$). The variability of this distribution is shown in Figure 2. Again the center is in close agreement with the theoretical variance of 25 (1/(0.2^2)). Finally, overlaying the histogram of distribution of averages with the empirical density curve and the normal density curve in Figure 3 shows that the distribution of averages is approximately normal.

Turning to the ToothGrowth dataset we begin with some exploratory analysis of the data. Figure 4 displays a histogram of tooth length split by supplement and dosage. The data here is quite variable with some supplements and dosage being quite clustered (vitamin C at 1.0mg for instance) while others are not (vitamin C at 2.0mg). To examine the efficacy of the treatments, each supplement is compared at the same dosage. That is to say the vitamin C and orange juice are compared at a dosage of 0.5mg and efficacy is determined via a independent t confidence interval taking VC - OJ. The same is carried out for the 1.0mg and 2.0mg dosages. The 0.5mg dosage yielded an interval of (-8.663, -1.837), the 1.0mg dosage (-9.003, -2.857) and finally the 2.0mg dosage (-3.477, 3.637). From these intervals we conclude that at 0.5mg and 1.0mg a supplement of vitamin C results in increased tooth growth relative to orange juice while at 2.0mg the treatment is inconclusive. To perform this analysis it was assumed the data was iid, that there were 60 seperate subjects to which the supplement and dosage combinations were administered to and that the variances were unequal.

Appendix

```
library(ggplot2)
library(grid)

set.seed(1)

Mean <- NULL

for(i in 1:1000)
     {
     Mean <- c(Mean, mean(rexp(40, 0.2)))
     }

qplot(Mean, geom = 'blank') + geom_histogram(alpha = 0.5) +
     geom_vline(xintercept = 1/0.2) + xlab("") +
     ggtitle("Averages of Exponential Distribution") +
     theme(plot.margin = unit(c(0,0,0,0), "cm"))
```

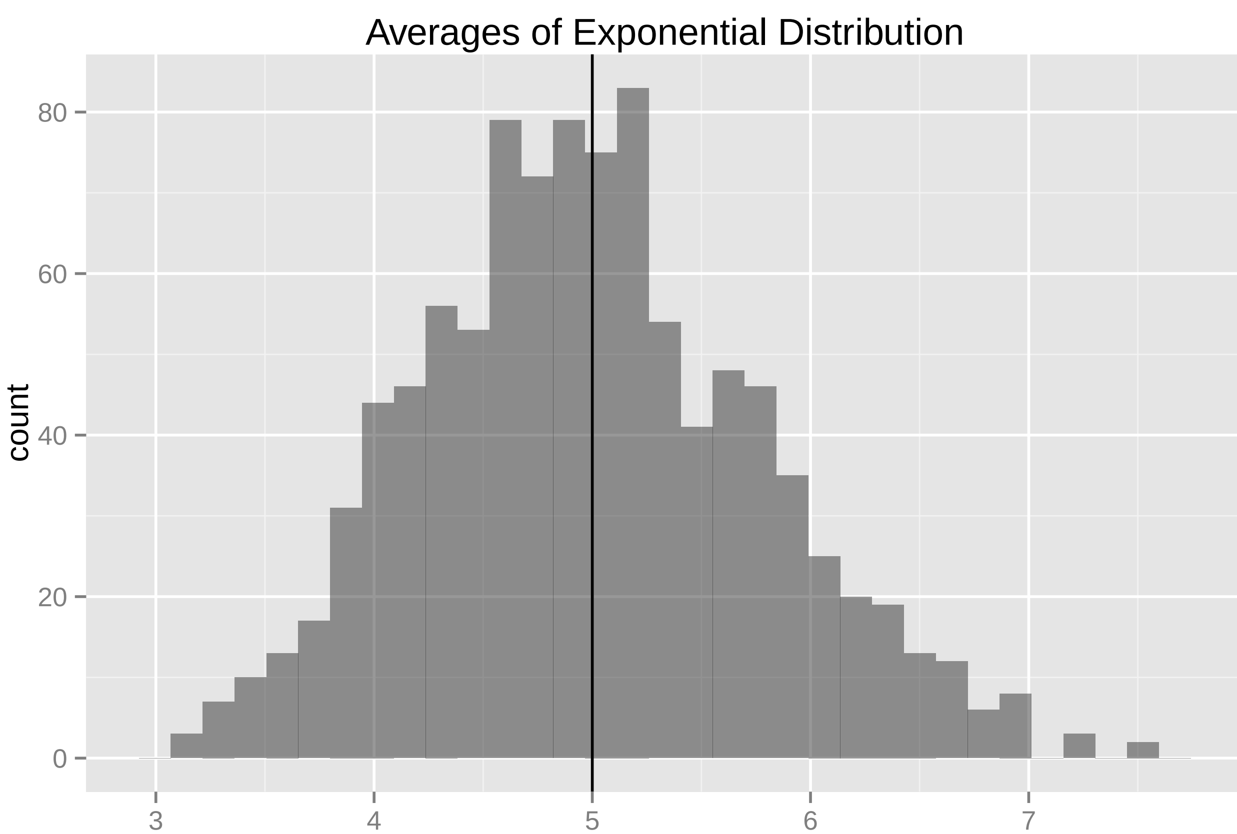

Figure 1

```
Variance <- NULL

for(i in 1:1000)
     {
     Variance <- c(Variance, var(rexp(40,0.2)))
     }

qplot(Variance, geom = 'blank') + geom_histogram(alpha = 0.5) +
     geom_vline(xintercept = 1/(0.2^2)) +
     xlab("") + ggtitle("Variance of Exponential Distribution") +
     theme(plot.margin = unit(c(0,0,0,0), "cm"))
```

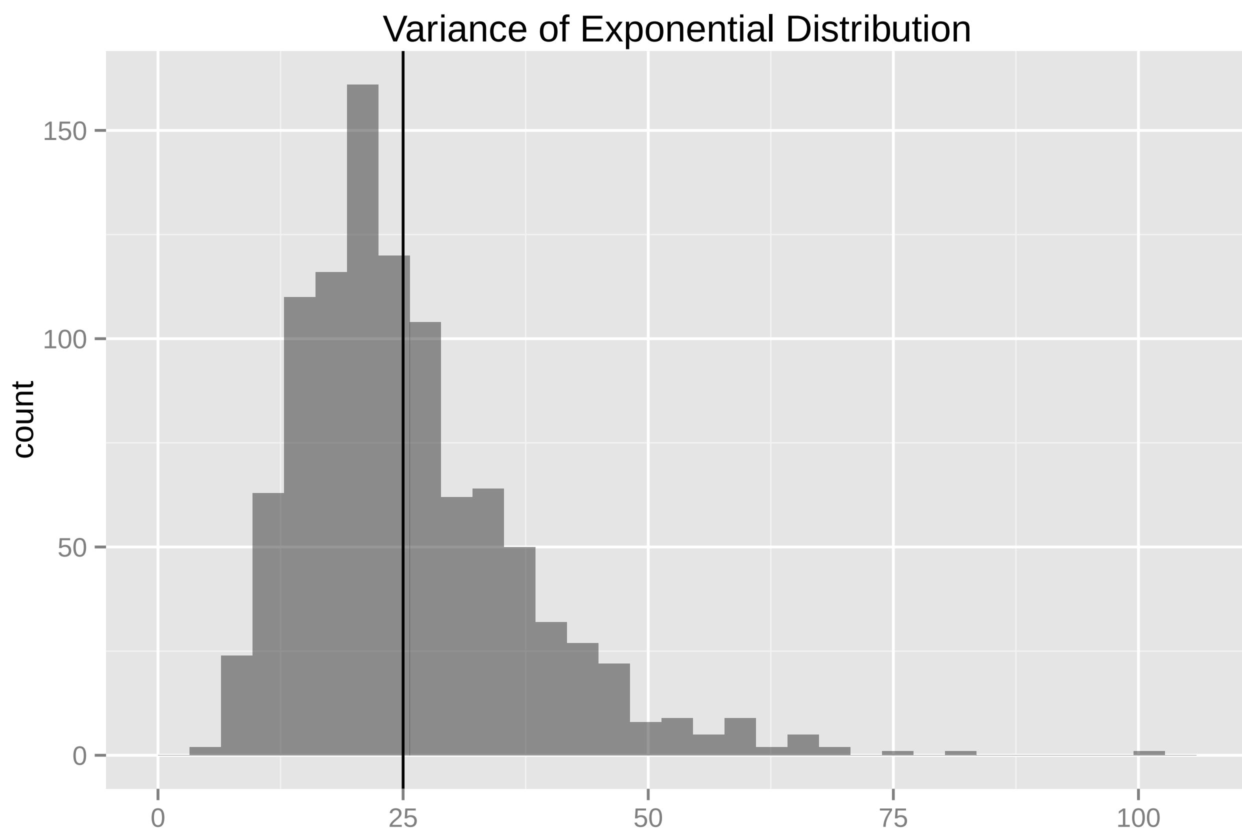

Figure 2

```
qplot(Mean, geom = 'blank') +
     geom_line(aes(y = ..density.., color = 'Empirical'), stat = 'density') +
     stat_function(fun = dnorm, aes(color = 'Normal'), args = list(mean = 1/0.2)) +
     geom_histogram(aes(y = ..density..), alpha = 0.5) +
     scale_color_manual(name = 'Density', values = c('red', 'blue')) + xlab("") +
     ggtitle("Distribution Curve of Averages of Exponential Distribution") +
     theme(plot.margin = unit(c(0,0,0,0), "cm"))
```

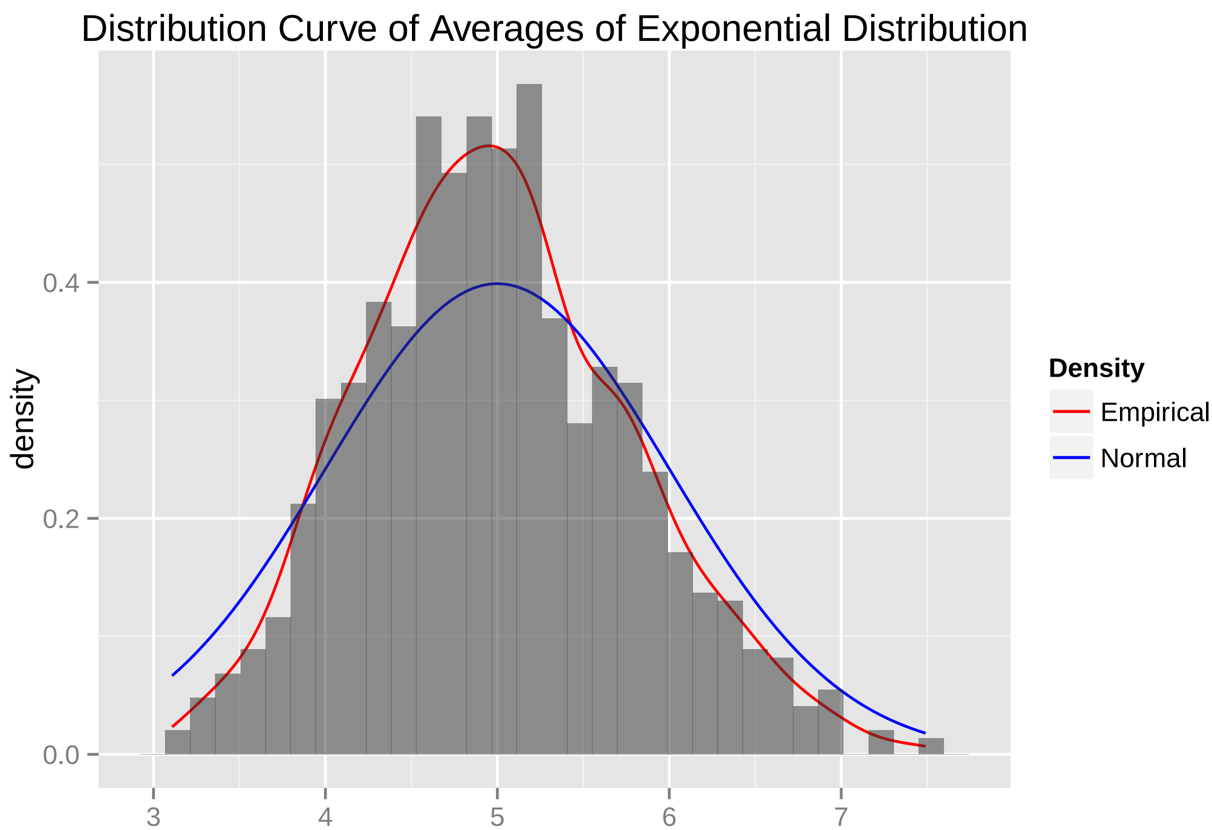

Figure 3

```r
library(datasets)

TG <- ToothGrowth
```

```r
ggplot(TG, aes(x = len), geom = 'blank') +
     geom_histogram(alpha = 0.5) + facet_grid(dose ~ supp) + xlab("") +
     ggtitle("Tooth Length by Supplement & Dosage") +
     theme(plot.margin = unit(c(0,0,0,0), "cm"))
```

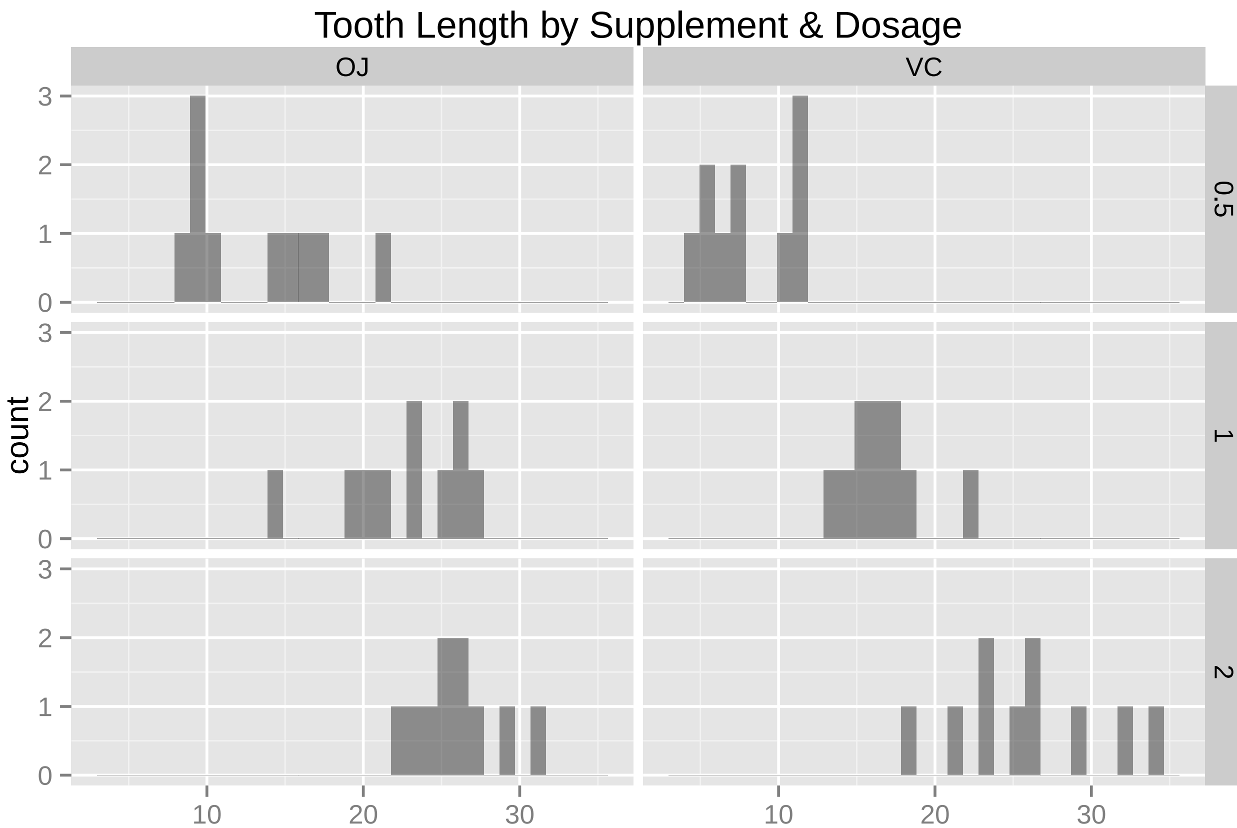

Figure 4

Function to evaluate indepedent t confidence interval.

```
IndeptConfInt <- function(x, y)
{
     #Calculates degrees of freedom
     df <- (var(x)/10 + var(y)/10)^2/((var(x)/10)/9 + (var(y)/10)/9)

     #Computes interval
     mean(x) - mean(y) + c(-1,1)*qt(0.975, df)*(var(x)/10 + var(y)/10)^(0.5)
}
```

Subsetting data.

```
VC0.5 <- TG[TG$supp == "VC" & TG$dose == "0.5", 1]
OJ0.5 <- TG[TG$supp == "OJ" & TG$dose == "0.5", 1]
VC1 <- TG[TG$supp == "VC" & TG$dose == "1", 1]
OJ1 <- TG[TG$supp == "OJ" & TG$dose == "1", 1]
VC2 <- TG[TG$supp == "VC" & TG$dose == "2", 1]
OJ2 <- TG[TG$supp == "OJ" & TG$dose == "2", 1]
```

Interval for each supplement at a dosage of 0.5mg.

```
IndeptConfInt(VC0.5, OJ0.5)
```

```
## [1] -8.663 -1.837
```

Interval for each supplement at a dosage of 1.0mg.

```
IndeptConfInt(VC1, OJ1)
```

```
## [1] -9.003 -2.857
```

Interval for each supplement at a dosage of 2.0mg.

```
IndeptConfInt(VC2, OJ2)
```

```
## [1] -3.477  3.637
```