Exponential Distribution compare to Central limit Theorem. Basic inferential data analysis

I D

```
# set seed for reproducability, set lambda to 0.2, 40 samples, 1000 simulations
set.seed(31)
lambda <- 0.2
n <- 40
simulations <- 1000
```

Load the ToothGrowth data and perform some basic exploratory data analyses

```
# load the data ToothGrowth
data(ToothGrowth)
# preview the structure of the data
str(ToothGrowth)
```

```
## 'data.frame':    60 obs. of  3 variables:
##  $ len : num  4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...
##  $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...
##  $ dose: num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

```
# preview first 5 rows of the data
head(ToothGrowth, 5)
```

```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
```

```
# Provide a basic summary of the data.
# data summary
summary(ToothGrowth)
```

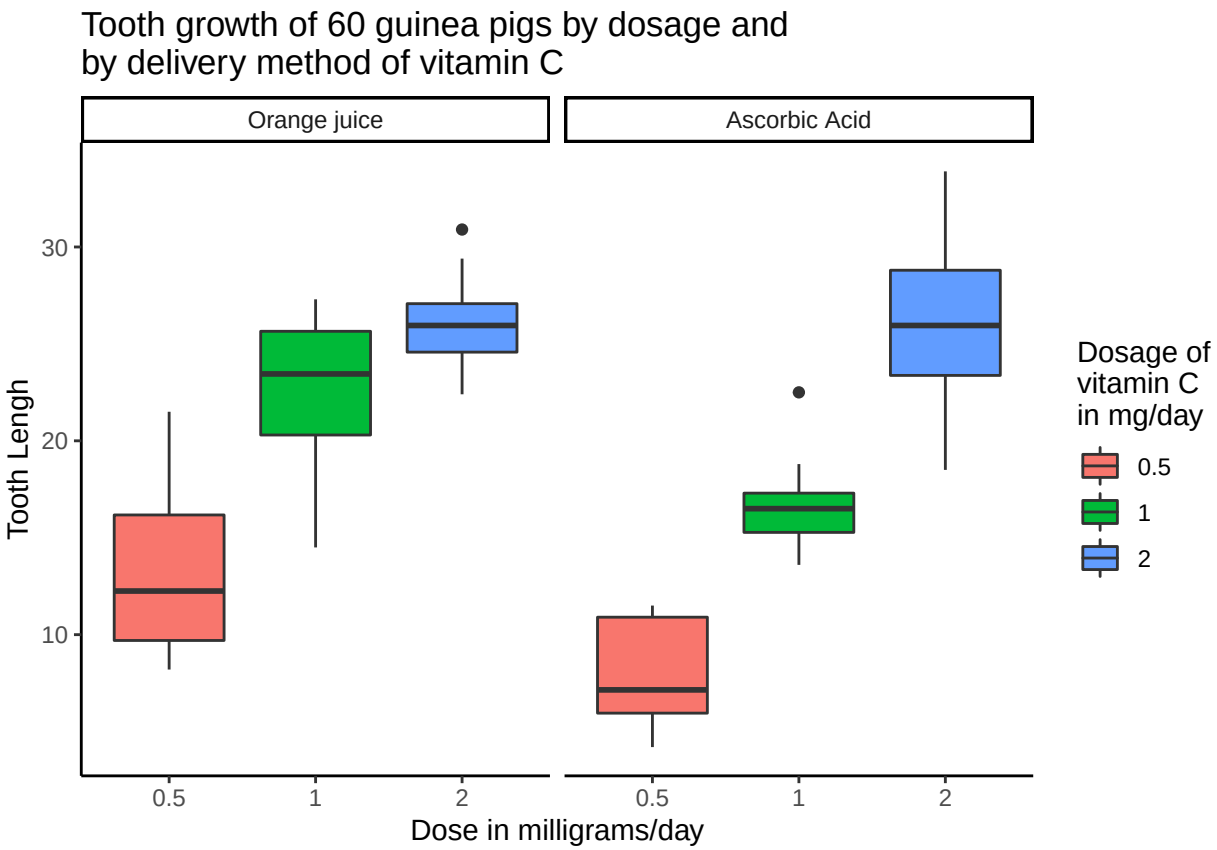
```
##       len        supp         dose
##  Min.   : 4.20   OJ:30   Min.   :0.500
##  1st Qu.:13.07   VC:30   1st Qu.:0.500
##  Median :19.25           Median :1.000
```

```
##  Mean   :18.81           Mean   :1.167
##  3rd Qu.:25.27           3rd Qu.:2.000
##  Max.   :33.90           Max.   :2.000
```

```r
# compare means of the different delivery methods
tapply(ToothGrowth$len,ToothGrowth$supp, mean)
```

```
##       OJ       VC
## 20.66333 16.96333
```

```r
# plot data graphically
library(ggplot2)
ggplot(ToothGrowth, aes(factor(dose), len, fill = factor(dose))) +
  geom_boxplot() +
  # facet_grid(.~supp)+
  facet_grid(.~supp, labeller = as_labeller(
    c("OJ" = "Orange juice",
      "VC" = "Ascorbic Acid"))) +
  labs(title = "Tooth growth of 60 guinea pigs by dosage and\nby delivery method of vitamin C",
       x = "Dose in milligrams/day",
       y = "Tooth Lengh") +
  scale_fill_discrete(name = "Dosage of\nvitamin C\nin mg/day") +
  theme_classic()
```



```
# Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose.
# comparison by delivery method for the same dosage
t05 <- t.test(len ~ supp,
              data = rbind(ToothGrowth[(ToothGrowth$dose == 0.5) &
                                       (ToothGrowth$supp == "OJ"),],
                           ToothGrowth[(ToothGrowth$dose == 0.5) &
                                       (ToothGrowth$supp == "VC"),]),
              var.equal = FALSE)
t1 <- t.test(len ~ supp,
             data = rbind(ToothGrowth[(ToothGrowth$dose == 1) &
                                      (ToothGrowth$supp == "OJ"),],
                          ToothGrowth[(ToothGrowth$dose == 1) &
                                      (ToothGrowth$supp == "VC"),]),
             var.equal = FALSE)
t2 <- t.test(len ~ supp,
             data = rbind(ToothGrowth[(ToothGrowth$dose == 2) &
                                      (ToothGrowth$supp == "OJ"),],
                          ToothGrowth[(ToothGrowth$dose == 2) &
                                      (ToothGrowth$supp == "VC"),]),
             var.equal = FALSE)
```

summary of the conducted t.tests, which compare the delivery methods by dosage,

take p-values and CI

```
summaryBYsupp <- data.frame(
  "p-value" = c(t05$p.value, t1$p.value, t2$p.value),
  "Conf.Low" = c(t05$conf.int[1],t1$conf.int[1], t2$conf.int[1]),
  "Conf.High" = c(t05$conf.int[2],t1$conf.int[2], t2$conf.int[2]),
  row.names = c("Dosage .05","Dosage 1","Dosage 2"))
```

```
# show data table
summaryBYsupp
```

```
##                p.value  Conf.Low Conf.High
## Dosage .05 0.006358607  1.719057  8.780943
## Dosage 1   0.001038376  2.802148  9.057852
## Dosage 2   0.963851589 -3.798070  3.638070
```

Conclusion

For dosage of .5 milligrams/day and 1 milligrams/day does matter the delivery method. the delivery method for 2 milligrams/day. For dosage of 2 milligrams/day the delivery method doesn't matter.