Inferential_Data_Analysis - The Effect of Vitamin C on Tooth Growth in Guinea Pigs

Banu Priya

8/14/2019

Overview

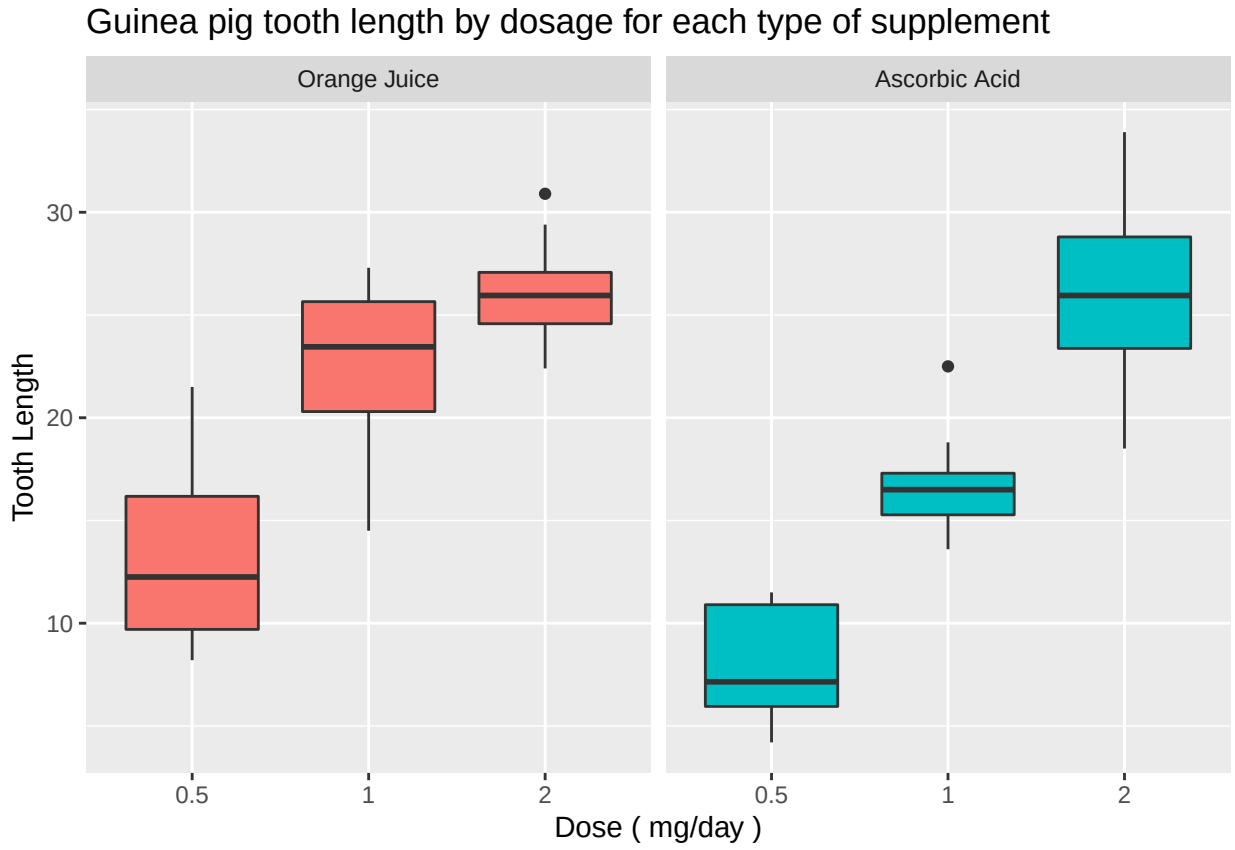
The purpose of this data analysis is to analyze the ToothGrowth data set by comparing the guinea tooth growth by supplement and dose. Below are the steps: - Exploratory data analysis on the given data set - To compare the analysis with the confidence intervals in order to make conclusions about the tooth growth

Load the ToothGrowth data & perform exploratory data analyses

```
library(datasets)
data(ToothGrowth)
str(ToothGrowth)
head(ToothGrowth)
summary(ToothGrowth)
```

```
library(ggplot2)
t = ToothGrowth
levels(t$supp) <- c("Orange Juice", "Ascorbic Acid")
ggplot(t, aes(x=factor(dose), y=len)) +
  facet_grid(.~supp) +
  geom_boxplot(aes(fill = supp), show_guide = FALSE) +
  labs(title="Guinea pig tooth length by dosage for each type of supplement",
    x="Dose ( mg/day )",
    y="Tooth Length")
```

```
## Warning: `show_guide` has been deprecated. Please use `show.legend`
## instead.
```



#Basic summary of the data The box plots indicates increasing the dosage increases the tooth growth. Orange juice is more effective than ascorbic acid for tooth growth when the dosage is .5 to 1.0 milligrams per day.

To summarize, Both the supplements are equally effective when the dosage is 2.0 milligrams per day.

#Using confidence intervals & hypothesis tests to compare tooth growth by supplement and dose #————————————Hypothesis #1————————————- Orange juice & ascorbic acid deliver the same tooth growth across the data set.

```
hypoth1<-t.test(len ~ supp, data = t)
hypoth1$conf.int
```

```
## [1] -0.1710156  7.5710156
## attr(,"conf.level")
## [1] 0.95
```

```
hypoth1$p.value
```

```
## [1] 0.06063451
```

The confidence intervals includes 0 and the p-value is greater than the threshold of 0.05. The null hypothesis cannot be rejected.

#————————————Hypothesis #2————————————- For the dosage of 0.5 mg/day, the two supplements deliver the same tooth growth.

```r
hypoth2<-t.test(len ~ supp, data = subset(t, dose == 0.5))
hypoth2$conf.int
```

```
## [1] 1.719057 8.780943
## attr(,"conf.level")
## [1] 0.95
```

```r
hypoth2$p.value
```

```
## [1] 0.006358607
```

The confidence interval does not include 0 and the p-value is below the 0.05 threshold. The null hypothesis can be rejected. The alternative hypothesis that 0.5 mg/day dosage of orange juice delivers more tooth growth than ascorbic acid is accepted.

#————————Hypothesis #3————————- For the dosage of 1 mg/day, the two supplements deliver the same tooth growth

```r
hypoth3<-t.test(len ~ supp, data = subset(t, dose == 1))
hypoth3$conf.int
```

```
## [1] 2.802148 9.057852
## attr(,"conf.level")
## [1] 0.95
```

```r
hypoth3$p.value
```

```
## [1] 0.001038376
```

The confidence interval does not include 0 and the p-value is smaller than the 0.05 threshold. The null hypothesis can be rejected. The alternative hypothesis that 1 mg/day dosage of orange juice delivers more tooth growth than ascorbic acid is accepted.

#————————Hypothesis #4————————- For the dosage of 2 mg/day, the two supplements deliver the same tooth growth

```r
hypoth4<-t.test(len ~ supp, data = subset(t, dose == 2))
hypoth4$conf.int
```

```
## [1] -3.79807  3.63807
## attr(,"conf.level")
## [1] 0.95
```

```r
hypoth4$p.value
```

```
## [1] 0.9638516
```

The confidence interval does include 0 and the p-value is larger than the 0.05 threshold. The null hypothesis cannot be rejected.

Conclusions and assumptions

1. Orange juice delivers more tooth growth than ascorbic acid for dosages 0.5 & 1.0.
2. Orange juice and ascorbic acid deliver the same amount of tooth growth for dose amount 2.0 mg/day.

Assumptions ->

- Normal distribution of the tooth lengths
- No other unmeasured factors are affecting tooth length