Hypothesis Testing

Yash Singhal

August 8, 2018

Tooth Growth Data

The response is the length of odontoblasts (cells responsible for tooth growth) in 60 guinea pigs. Each animal received one of three dose levels of vitamin C (0.5, 1, and 2 mg/day) by one of two delivery methods, orange juice or ascorbic acid (a form of vitamin C and coded as VC).

```
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 ...
```

The dataset consist of 60 observations, They are adminstrated different level of dosage with different explore methods Lets look into it more closely

```
df <- ToothGrowth
table(df$supp,df$dose)
```

```
##
##      0.5  1  2
##   OJ  10 10 10
##   VC  10 10 10
```

The dataset is well consistent with 10 observation of tooth length for each of the differnt supplement type and dosage

Exploratary Data Analysis

- There Seems to be some relationship Between the Tooth Growth and Dosage , with Higher Dosage leading to Higher Tooth Growth (Figure 1)
- The Relationship here between the Dosage and Delivery Method isn't quite Strong (Figure 2)

Hypothesis Testing

We want to Check for the Hypothesis that

- There is a substantial Increase in the Tooth Growth when adminstrated Different Doses
- Tooth Growth does not depend on the Delivery Method

Asumptions

- The Tooth Growth resembles normally distributed
- There would be equal Variances between dfferrent pairs of observations as these observation are from the same population distrbution

Part-1

(H0 -> There is no difference in Tooth Length when adminstraed a dosage of 2 and 1) (H1 -> There is a Substantial Difference in Tooth Length when adminstrated a dosage of 2 and 1)

```
filter(df,((dose==1) | (dose==2))) -> experiment
t.test(experiment$len~as.factor(experiment$dose),var=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  experiment$len by as.factor(experiment$dose)
## t = -4.9005, df = 38, p-value = 1.811e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.994387 -3.735613
## sample estimates:
## mean in group 1 mean in group 2
##          19.735          26.100
```

A small P value well below the signifciance level of 0.5 indicates a substantial difference and null hypotheis can be discarded in favour of alternative

Part-2

(H0 -> There is no difference in Tooth Length when adminstraed a dosage of 1 and 0.5) (H1 -> There is a Substantial Difference in Tooth Length when adminstrated a dosage of 1 and 0.5)

```
filter(df,((dose==1) | (dose==0.5))) -> experiment
t.test(experiment$len~as.factor(experiment$dose),var=TRUE)
```

```
##
##  Two Sample t-test
##
## data:  experiment$len by as.factor(experiment$dose)
## t = -6.4766, df = 38, p-value = 1.266e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.983748  -6.276252
## sample estimates:
## mean in group 0.5   mean in group 1
##            10.605            19.735
```

Here the Same Low p value indicates there indeed is a strong relationship between the tooth length and whether dose is adminstrated as 1mg or 0mg

Part-3

(H0 -> There is no difference in Tooth Length when adminstraed a injection using VJ Or VC supplement) (H1 -> There is a Substantial Difference in Tooth Length when adminstrated a injection using VJ or VC Supplement)

```
t.test(df$len~df$supp,var=TRUE)
```

```
##
##  Two Sample t-test
```

```
##
## data:  df$len by df$supp
## t = 1.9153, df = 58, p-value = 0.06039
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1670064  7.5670064
## sample estimates:
## mean in group OJ mean in group VC
##         20.66333         16.96333
```

A p value of 0.6 is not below the significant value and hence we cannot reject null in favour of alternative

Conclusions

- From Part 1, Part 2 and Figure 1 and Figure 2 we can conclude that the toothlength indeed differs depended on the dose adminstrated
- We cannot stastically conclude that the tooth length differs depending on the delievery method as illustrated in Part 3 of hypothesis Testing

Appendix

```
ggplot(aes(x=as.factor(dose),y=len),data=df)+geom_boxplot()+labs(x="Dosage")
```

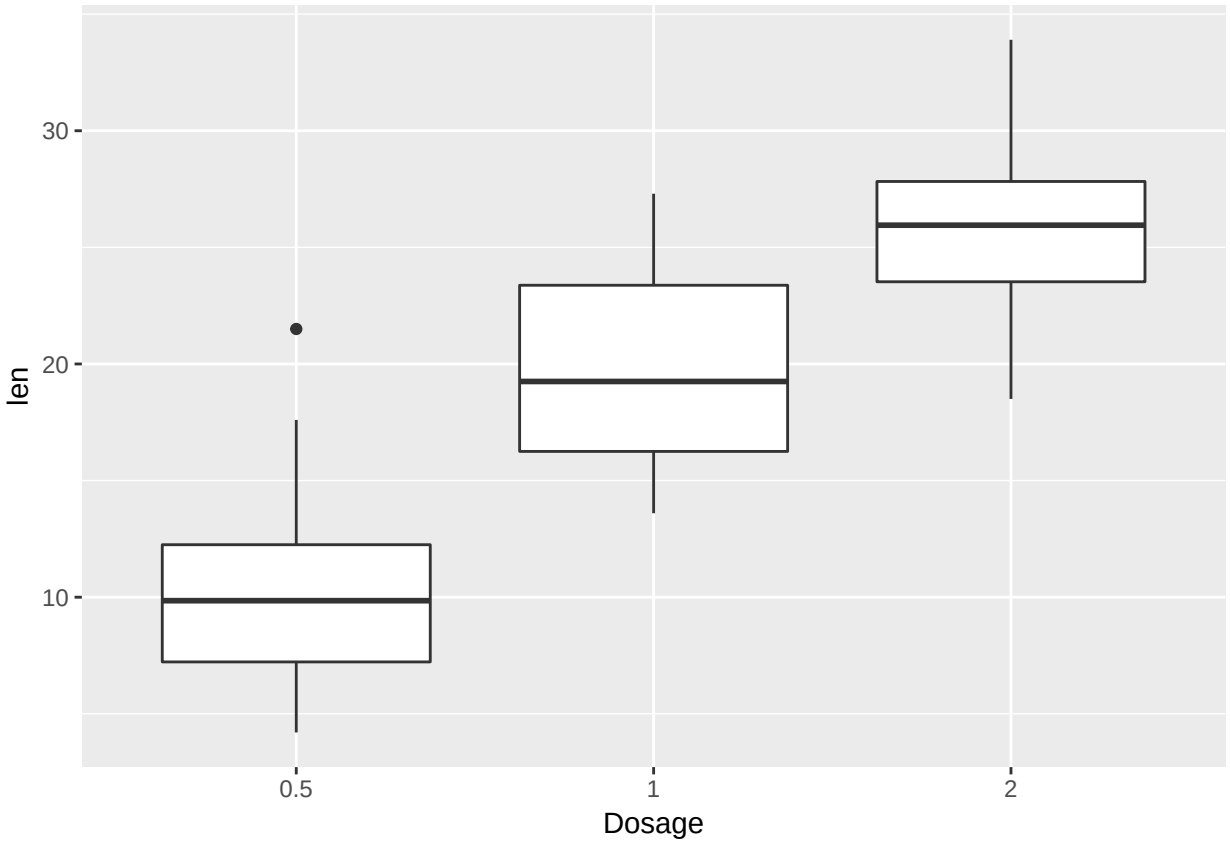

Figure 1 Boxplot of tooth length vs dosage

```
ggplot(aes(x=as.factor(supp),y=len),data=df)+geom_boxplot()+labs(x="Dosage")
```

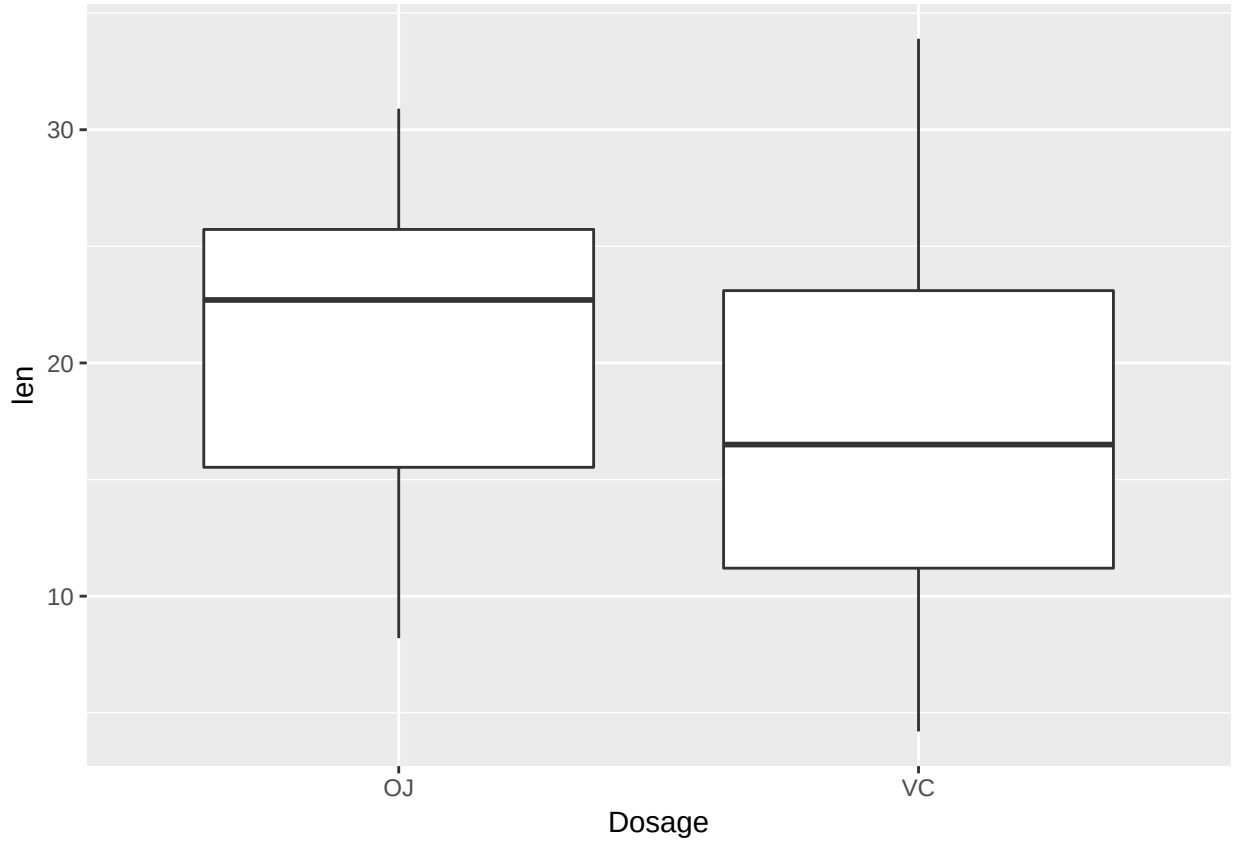

Figure 2 Boxplot of tooth length vs injection method