

## Statistical Inference Course Project 2:

### analysis of tooth growth for guinea pigs by supp and dose

by Henrik Gjerner, 27th September 2015

This is the project for the statistical inference class. The data analysis will be in four parts: 1. Exploratory analysis of the data 2. Basic data summary 3. Comparison of supp and dose 4. Conclusions and assumptions. The analysis will compare samples to theoretical distributions to evaluate their similarities based on the parameters below.

#### 1. Load the ToothGrowth data and perform some basic exploratory data analyses

```
library(datasets)
library(ggplot2)

ggplot(ToothGrowth, aes(x=dose, y=len)) + ggtitle("ToothGrowth Length vs Dose by Delivery Method") + geom_boxplot(aes(fill=supp))
```



The plot above shows 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)).

## 2. Provide a basic summary of the data.

```
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
```

## 3. Use confidence intervals and/or hypothesis tests to compare tooth growth by supp and dose

The hypothesis for the following tests will be:

- H0: There is no difference between orange juice and ascorbic acid for a given dose
- H1: There is a difference between orange juice and ascorbic acid for a given dose

We are assuming the sample components are independent and identically distributed. `# {r confidence}`  
`# t.test(len ~ supp, data=ToothGrowth)` # Across all three dosages the 95% confidence interval does contain 0, so we cannot reject the null hypothesis so the data suggests that there is no statistical significance between OJ and VC for all doses.

Dividing the data into the three subgroups by dose we then start by analysing the 0.5 dose for OJ and VC

```
Dose0.5 = ToothGrowth[ToothGrowth$dose == 0.5, ]
t.test(len ~ supp, data=Dose0.5)
```

```
##
## Welch Two Sample t-test
##
## data: len by supp
## t = 3.1697, df = 14.969, p-value = 0.006359
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.719057 8.780943
## sample estimates:
## mean in group OJ mean in group VC
##           13.23           7.98
```

The 95% confidence interval does not contain 0, so we should reject the null hypothesis so the data suggests that the OJ has higher statistical significance than VC given at the 0.5 dose level.

The second group contains the 1.0 dose for OJ and VC

```
Dose1 = ToothGrowth[ToothGrowth$dose == 1, ]
t.test(len ~ supp, data=Dose1)
```

```
##
## Welch Two Sample t-test
##
## data: len by supp
## t = 4.0328, df = 15.358, p-value = 0.001038
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 2.802148 9.057852
## sample estimates:
## mean in group OJ mean in group VC
## 22.70 16.77
```

The 95% confidence interval does not contain 0, so we should reject the null hypothesis. Again the data suggests that the OJ has higher statistical significance than VC given at the 1.0 dose level.

The third group contains the 2.0 dose for OJ and VC

```
Dose2 = ToothGrowth[ToothGrowth$dose == 2, ]
t.test(len ~ supp, data=Dose2)
```

```
##
## Welch Two Sample t-test
##
## data: len by supp
## t = -0.046136, df = 14.04, p-value = 0.9639
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -3.79807 3.63807
## sample estimates:
## mean in group OJ mean in group VC
## 26.06 26.14
```

The 95% confidence interval does contain 0, so we cannot reject the null hypothesis. So data suggests that there is no statistical significance between OJ and VC at the 2.0 dose level.

## 4.State your conclusions and the assumptions needed for your conclusions.

In the previous part we did a number of tests comparing different doses of Orange juice to ascorbic acid and we can conclude that for smaller doses feed into guinea pigs, Orange Juice performs better so:

- Using 0.5-1.0 doses of Orange Juice gives a significant better teeth growth than using similar doses of Ascorbic Acid
- However this is not the case when using a higher dose of 2.0