Test of the effect of vitamin C on tooth growth for guinea pigs

Nadia Chylak

As part of this analysis, we will look into the effect of vitamin C on tooth growth for guinea pigs. In particular, we will examine whether the dosage and mode of administration of vitamin C have any effect on the measured tooth growth of our sample in order to infer conclusions about the wider population.

Basic Data Exploration

```
library(datasets)
data("ToothGrowth")
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 ...
```

```
table(ToothGrowth$supp)
```

```
##
## OJ VC
## 30 30
```

```
table(ToothGrowth$dose)
```

```
##
## 0.5   1   2
##  20  20  20
```

ToothGrowth is a dataframe of 60 observations on 3 variables recording the effect of vitamin C on tooth growth for guinea pigs. The three variables represent:

- `len` : tooth length (*numeric*)
- `supp` : supplement type, either OJ for orange juice or VC for ascorbic acid (*factor*)
- `dose` : dose received in mg/day (*numeric*)

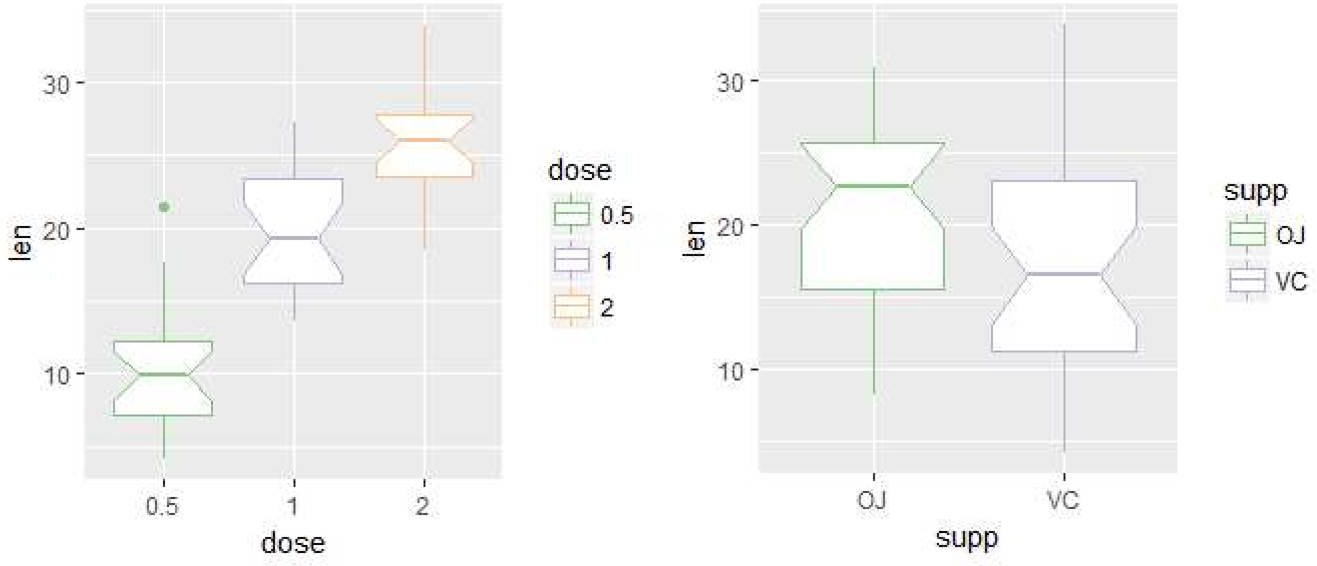
Each 60 guinea pigs received one of three dose levels of vitamin C (0.5, 1 or 2 mg/day) by one of two delivery methods, consequently to which the length of their odontoblasts (cells responsible for tooth growth) were measured.

The subjects are divided into six groups:

Number of subjects	**OJ**	**VC**	**Total**
0.5	10	10	20

Number of subjects	OJ	VC	Total
1	10	10	20
2	10	10	20
Total	30	30	60

Figure 1: Visualisation of tooth length in terms of dosage and supplement types.



Based on this initial exploration, it looks like higher doses of vitamin C have for effect higher tooth length. Also, orange juice seems to have a greater effect on increasing tooth length than ascorbic acid. We will now test the robustness of these assumptions.

Hypothesis testing

Different vitamin C dosage have the same effect on tooth growth

As part of our t-test, we will use the following as hypothesis:

- H_0: Different vitamin C dosage have the same effect on tooth growth, i.e. the average tooth growth for subjects receiving a dose of 0.5 mg/day (μ_1) is the same as for subjects receiving a dose of 2 mg/day (μ_2), which is the same as testing whether $\mu_1 - \mu_2 = 0$
- H_a: Vitamin C dosage has an effect on tooth growth, i.e. $\mu_1 - \mu_2 \neq 0$.

We will build a confidence interval at a level of 95%. If the confidence interval does not contain the desired value of the parameter (in our case zero), we can exclude H_0 with 95% confidence.

For our t-test, we need to make the following assumptions:

- Tooth length is approximately normally distributed within the population of guinea pigs.
- Our sample is a simple random sample, i.e. the data is collected from a randomly selected portion of the total population of guinea pigs
- The variance of tooth length is equal for subjects receiving a dose of 0.5 mg/day and subjects receiving a dose of 2 mg/day.

```
doseMin <- ToothGrowth$len[ToothGrowth$dose == 0.5]
doseMax <- ToothGrowth$len[ToothGrowth$dose == 2]
t.test(doseMin, doseMax, paired = FALSE , var.equal = TRUE, conf.level = 0.95)$conf.int[1:2]
```

```
## [1] -18.15352 -12.83648
```

For the sake of learning, we reproduce this confidence interval manually.

```
n <- length(doseMin) # doseMax is of the same length
s <- sqrt(((n-1)*var(doseMin) + (n-1)*var(doseMax))/(2*n-2))
mean(doseMin) - mean(doseMax) + c(-1,1) * qt(0.975, df = 2*n-2) * s * sqrt(1/n+1/n)
```

```
## [1] -18.15352 -12.83648
```

The confidence interval does not contain zero and so this allows us to invalidate hypothesis H_0.

Both supplement types have the same effect on tooth growth

As part of our t-test, we will use the following as hypothesis:

- H_0: Both supplement types have the same effect on tooth growth, i.e. the average tooth growth for subjects receiving vitamin C in the form of OJ (μ_1) is the same as for subjects receiving vitamin C in the form of VC (μ_2), which is the same as testing whether $\mu_1 - \mu_2 = 0$
- H_a: Vitamin C supplement types have an effect on tooth growth, i.e. $\mu_1 - \mu_2 \neq 0$.

We will build a confidence interval at a level of 95%. If the confidence interval does not contain the desired value of the parameter (in our case zero), we can exclude H_0 with 95% confidence.

For our t-test, we need to make the following assumptions:

- Tooth length is approximately normally distributed within the population of guinea pigs.
- Our sample is a simple random sample, i.e. the data is collected from a randomly selected portion of the total population of guinea pigs
- The variance of tooth length is equal for subjects receiving vitamin C in the form of OJ and subjects receiving vitamin C in the form of VC.

```
suppOJ <- ToothGrowth$len[ToothGrowth$supp == "OJ"]
suppVC <- ToothGrowth$len[ToothGrowth$supp == "VC"]
t.test(suppOJ, suppVC, paired = FALSE , var.equal = TRUE, conf.level = 0.95)$conf.int[1:2]
```

```
## [1] -0.1670064  7.5670064
```

For the sake of learning, we reproduce this confidence interval manually.

```
n <- length(suppOJ) # suppVC is of the same length
s <- sqrt(((n-1)*var(suppOJ) + (n-1)*var(suppVC))/(2*n-2))
mean(suppOJ) - mean(suppVC) + c(-1,1) * qt(0.975, df = 2*n-2) * s * sqrt(1/n+1/n)
```

```
## [1] -0.1670064  7.5670064
```

The confidence interval contains zero and so we can not exclude that H_0 is valid.

Appendix

```r
library(ggplot2)

# Plot 1
ToothGrowth <- transform(ToothGrowth, dose = factor(dose))
g1 <- ggplot(ToothGrowth, aes(x=dose, y=len, color=dose))
g1 <- g1 + geom_boxplot(notch=TRUE)
g1 <- g1 + scale_color_brewer(palette = "Accent")

# Plot 2
g2 <- ggplot(ToothGrowth, aes(x=supp, y=len, color=supp))
g2 <- g2 + geom_boxplot(notch=TRUE)
g2 <- g2 + scale_color_brewer(palette = "Accent")

# Arrange
library(gridExtra)
grid.arrange(g1, g2, nrow = 1)
```