Statistical Inference - Course Project pt 2

Richard Hardy

2019-12-15

Part 2

Odontoblast growth in 60 guinea-pigs.

Vars:

- Odontoblast length (units not found),
- Vitamin C dose (mg/day),
- Method of delivery (Orange Juice or as ascorbic acid)

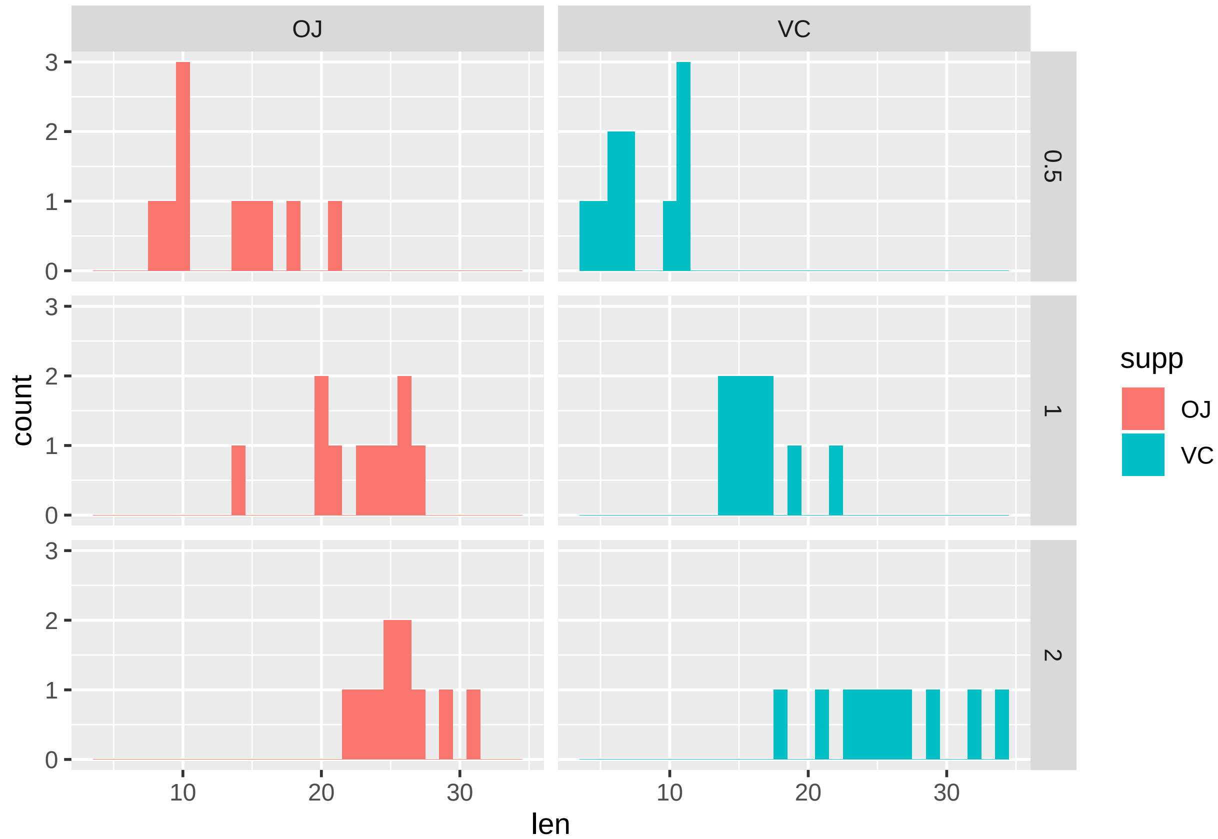
See link: dataset information

Exploratory Analysis

```
library(datasets)
library(ggplot2)
data(ToothGrowth)
```

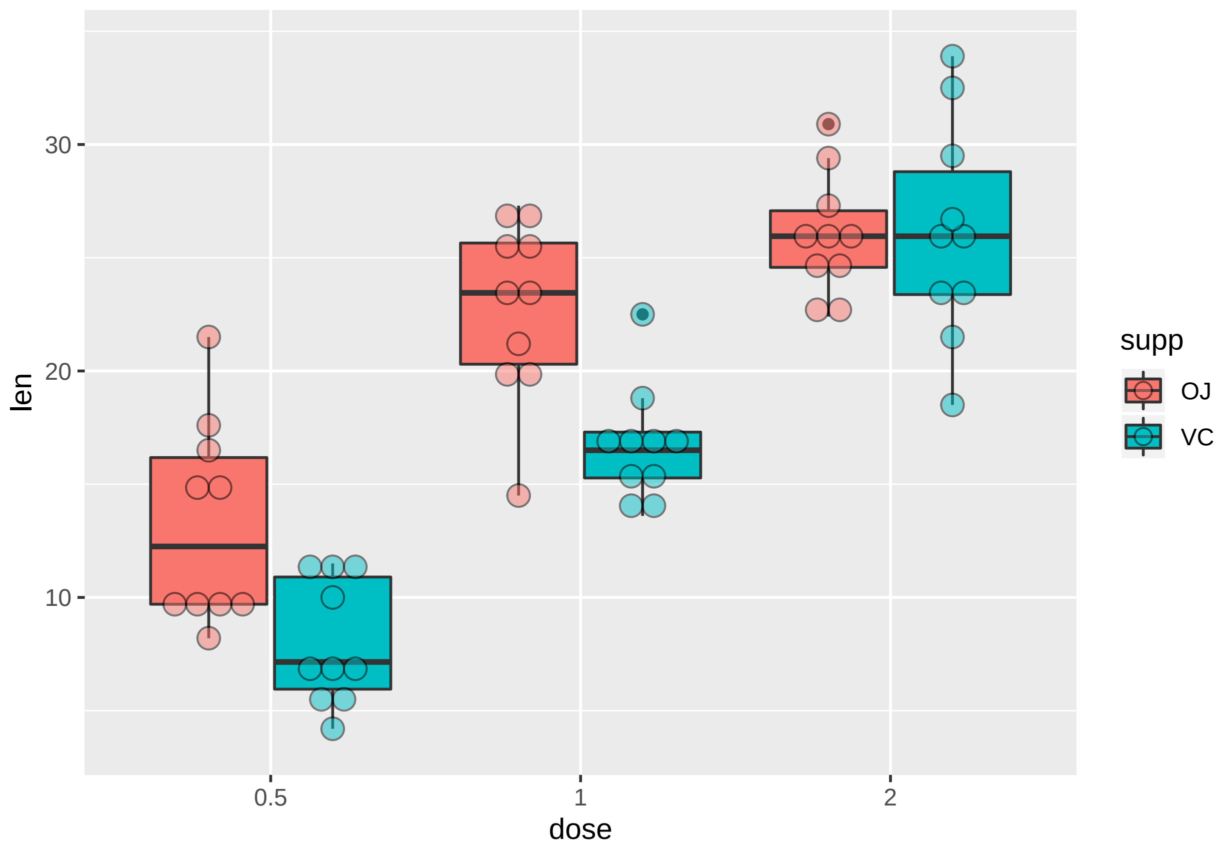
```
#plot 1: histograms showing frequency distributions:

g <- ggplot(ToothGrowth, aes(x=len,fill=supp))
g <- g + geom_histogram(binwidth=1)
g <- g + facet_grid(dose~supp)
g
```




```r
#plot 2: more detail: dotplots/boxplots shown distributions and means:

ToothGrowth$dose <- as.factor(ToothGrowth$dose)
h <- ggplot(ToothGrowth, aes(x = dose, y = len, fill = supp))
h <- h + geom_boxplot(position = position_dodge(0.8))
h <- h + geom_dotplot(binaxis ='y', stackdir = 'center', binwidth = 1,
                      position = position_dodge(0.8), alpha=0.5)
h
```



From the above plot, there appears to be a dose effect, and delivery type may have an effect at dose levels 0.5 mg/day and 1 mg/day, but not at dose level 2 mg/day.

```
# Create vectors for summary statistics and later t-tests:
OJgroup.5 <- ToothGrowth$len[ToothGrowth$supp == "OJ" & ToothGrowth$dose == 0.5]
OJgroup1 <- ToothGrowth$len[ToothGrowth$supp == "OJ" & ToothGrowth$dose == 1]
OJgroup2 <- ToothGrowth$len[ToothGrowth$supp == "OJ" & ToothGrowth$dose == 2]

VCgroup.5 <- ToothGrowth$len[ToothGrowth$supp == "VC" & ToothGrowth$dose == 0.5]
VCgroup1 <- ToothGrowth$len[ToothGrowth$supp == "VC" & ToothGrowth$dose == 1]
VCgroup2 <- ToothGrowth$len[ToothGrowth$supp == "VC" & ToothGrowth$dose == 2]

# Summary statistics - means (on plot, viewed left to right):

mns <- c(mean(OJgroup.5), mean(VCgroup.5), mean(OJgroup1), mean(VCgroup1),
         mean(OJgroup2), mean(VCgroup2))

mns
```

```
## [1] 13.23  7.98 22.70 16.77 26.06 26.14
```

```
# Summary statistics - standard errors of the mean (on plot, viewed left to right):

ses <- c(sd(OJgroup.5), sd(VCgroup.5), sd(OJgroup1), sd(VCgroup1),
         sd(OJgroup2), sd(VCgroup2))
```

```r
ses
```

```
## [1] 4.459709 2.746634 3.910953 2.515309 2.655058 4.797731
```

Analysis: Independent t-tests - unequal variances

Unequal variances are used given the distributions shown on plots 1 and 2, and that there is no reason (that I can tell) to assume that they should be equal between dose / delivery groups. The data are not paired, assumed to be randomised to 6 groups of 10 from N = 60.

Set:

- μ_1 = mean of `VCgroup.5`
- μ_2 = mean of `VCgroup1`
- μ_3 = mean of `VCgroup2`
- μ_4 = mean of `OJgroup.5`
- μ_5 = mean of `OJgroup1`
- μ_6 = mean of `OJgroup2`

Supp types hypotheses:

- $H1_0 : \mu_1 = \mu_4$ and $H1_a : \mu_1 \neq \mu_4$
- $H2_0 : \mu_2 = \mu_5$ and $H2_a : \mu_2 \neq \mu_5$
- $H3_0 : \mu_3 = \mu_6$ and $H3_a : \mu_3 \neq \mu_6$

Dosage hypotheses : VC group

- $H4_0 : \mu_1 = \mu_2$ and $H4_a : \mu_1 \neq \mu_2$
- $H5_0 : \mu_1 = \mu_3$ and $H5_a : \mu_1 \neq \mu_3$
- $H6_0 : \mu_2 = \mu_3$ and $H6_a : \mu_2 \neq \mu_3$

Dosage hypotheses : OJ group

- $H7_0 : \mu_4 = \mu_5$ and $H7_a : \mu_4 \neq \mu_5$
- $H8_0 : \mu_4 = \mu_6$ and $H8_a : \mu_4 \neq \mu_6$
- $H9_0 : \mu_5 = \mu_6$ and $H9_a : \mu_5 \neq \mu_6$

```r
# Is there a difference in odontoblast length between supplementation types?

# At dose = 0.5
# H1: VC vs OJ
H1 <- t.test(OJgroup.5, VCgroup.5, paired=FALSE, var.equal=FALSE)

# At dose = 1
# H2: VC vs OJ
H2 <- t.test(OJgroup1, VCgroup1, paired=FALSE, var.equal=FALSE)

# At dose = 2
# H3: VC vs OJ
H3 <- t.test(OJgroup2, VCgroup2, paired=FALSE, var.equal=FALSE)
```

```r
#Is there a within-group difference in odontoblast length between dosages?

# In the VC group:

# H4: dose = 0.5 vs dose = 1
H4 <- t.test(VCgroup.5, VCgroup1, paired=FALSE, var.equal=FALSE)
# H5: dose = 0.5 vs dose = 2
H5 <- t.test(VCgroup.5, VCgroup2, paired=FALSE, var.equal=FALSE)
# H6: dose = 1 vs dose = 2
H6 <- t.test(VCgroup1, VCgroup2, paired=FALSE, var.equal=FALSE)

# In the OJ group:

# H7: 0.5 vs 1
H7 <- t.test(OJgroup.5, OJgroup1, paired=FALSE, var.equal=FALSE)
# H8: 0.5 vs 2
H8 <- t.test(OJgroup.5, OJgroup2, paired=FALSE, var.equal=FALSE)
# H9: 1 vs 2
H9 <- t.test(OJgroup1, OJgroup2, paired=FALSE, var.equal=FALSE)
```

With 9 separate independent t-tests for significance, conclusions are based on p-values using bonferroni correction for multiple testing with $m = 9$.

```r
bonferroni <- .05 / 9

pvals = c(H1[[3]], H2[[3]], H3[[3]], H4[[3]], H5[[3]], H6[[3]], H7[[3]], H8[[3]], H9[[3]])

lt.bonf <- p.adjust(pvals, method = "bonferroni") < 0.05

ests <- cbind(pvals, lt.bonf)

# For each of the nine hypotheses, the following indicates whether p values are
# below the bonferroni-adjusted threshold
ests
```

```
##               pvals lt.bonf
##  [1,] 6.358607e-03       0
##  [2,] 1.038376e-03       1
##  [3,] 9.638516e-01       0
##  [4,] 6.811018e-07       1
##  [5,] 4.681577e-08       1
##  [6,] 9.155603e-05       1
##  [7,] 8.784919e-05       1
##  [8,] 1.323784e-06       1
##  [9,] 3.919514e-02       0
```

Assumptions:

- Unequal population variances - i.e. tests do not assume equal population variances for each combination of dose / supp.
- There is sufficient power to detect clinically relevant differences in tooth length with $n = 10$ in each group. Power calculations were not performed.

- Guinea-pigs were randomised to each of the groups, and this adequately controls for unmeasured confounding factors unrelated to dose or supplement type.
- A random draw of odontoblast length x follows a pdf that is t-distributed in each group.

Conclusions:

There is evidence for a dose-dependent effect of vitamin C on odontoblast growth, if delivered as ascorbic acid, up to at least 2.0 mg/day.

- data supports a dose-dependent effect of vitamin C on mean odontoblast length, at all tested dosage levels, if delivered as ascorbic acid (reject $H4_0$, $H5_0$, $H6_0$ at bonferroni $p < 0.0056$).

There is evidence for a higher effect of orange juice delivered vitamin C than ascorbic acid on odontoblast length, however this difference appears only at dosages of 1.0 mg/day.

- data supports a difference in mean odontoblast length between VC and OJ delivery methods, for a dose of 1 mg/day (reject $H2_0$ at bonferroni $p < 0.0056$).
- data does not support a difference in mean odontoblast length between VC and OJ vitamin C delivery methods, at a dose of 0.5 mg/day (fail to reject $H1_0$ at bonferroni $p < 0.0056$).
- data does not support a difference in mean odontoblast length between VC and OJ vitamin C delivery methods, at a dose of 2.0 mg/day (fail to reject $H3_0$ at bonferroni $p < 0.0056$).

Increasing the dosage of vitamin C delivered via orange juice above 0.5 mg/day appears to increase odontoblast length, however this effect attenuates above a dose of 1 mg/day

- data supports a dose-related effect of vitamin C on mean odontoblast length above 0.5 mg/day if delivered via orange juice (reject $H7_0$ and $H8_0$ at bonferroni $p < 0.0056$).
- data does not support a difference in mean odontoblast length between doses of 1 and 2 mg/day of Vitamin C if delivered via orange juice (fail to reject $H9_0$ at bonferroni $p < 0.0056$).

Following the conclusions of this analysis, the data suggests a further hypothesis that additional factors present in orange juice may have a cumulative effect with vitamin C, on odontoblast length. This may be a question for further analysis