Statistical Inference Project-2

Adekunle Adeniran

2023-11-10

PART 2 Basic Inferential Data Analysis

Loading the required libraries and take a peek at the dataset

```
library(stats)
library(ggplot2)
library(kableExtra)
data("ToothGrowth")
head(ToothGrowth)
```

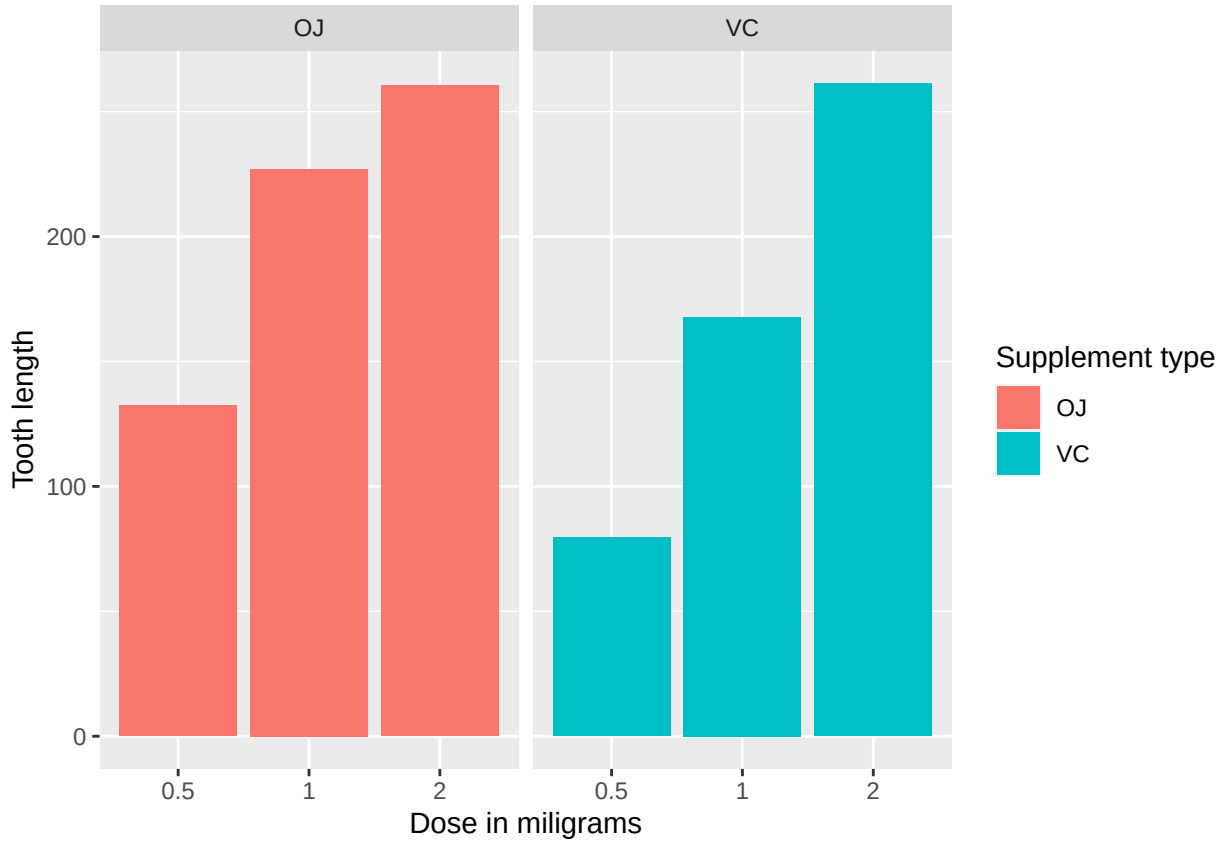
```
##    len supp dose
## 1  4.2   VC  0.5
## 2 11.5   VC  0.5
## 3  7.3   VC  0.5
## 4  5.8   VC  0.5
## 5  6.4   VC  0.5
## 6 10.0   VC  0.5
```

Summary of 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
```

```
ggplot(data=ToothGrowth, aes(x=as.factor(dose), y=len, fill=supp)) +
    geom_bar(stat="identity",) +
    facet_grid(. ~ supp) +
    xlab("Dose in miligrams") +
    ylab("Tooth length") +
    guides(fill=guide_legend(title="Supplement type"))
```



Using Confidence Intervals to compare tooth growth by supp and dose.

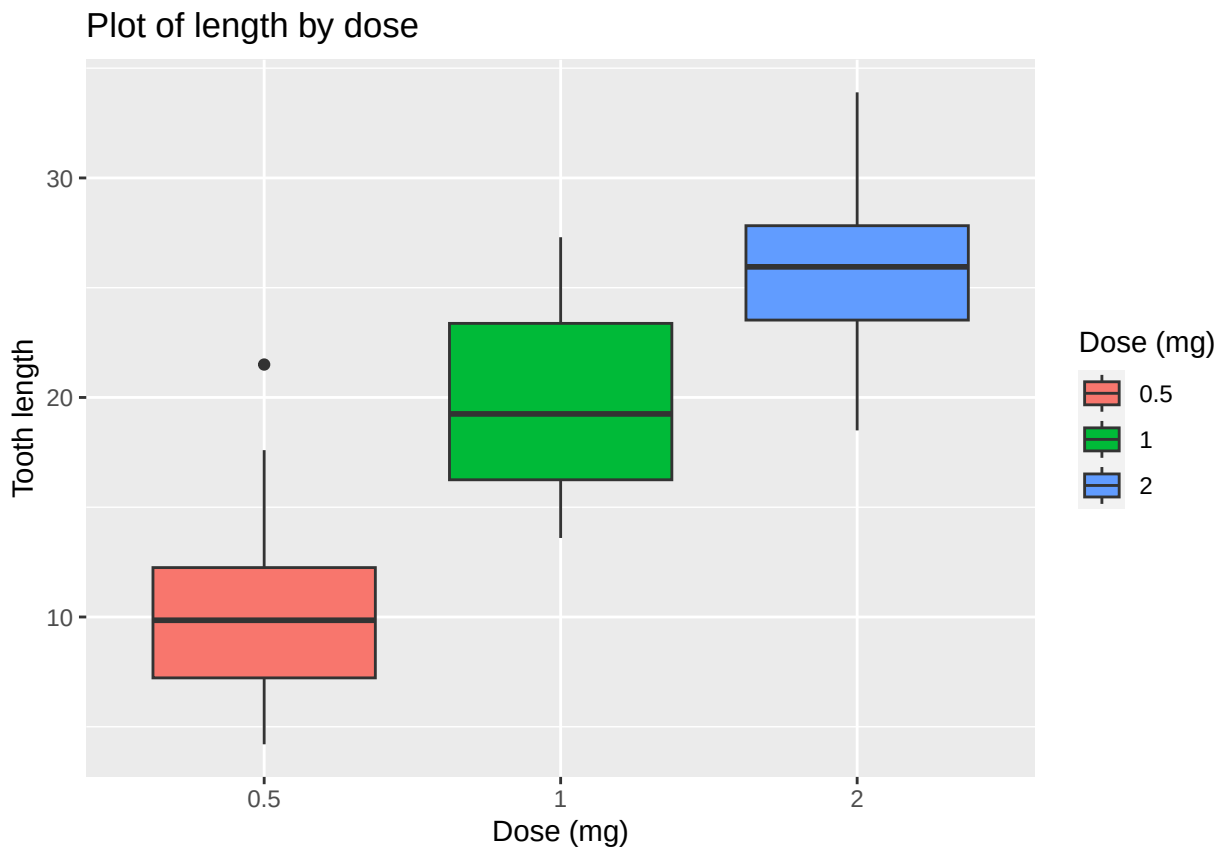
```
unique(ToothGrowth$dose)
```

```
## [1] 0.5 1.0 2.0
```

There are 3 dose groups: 0.5, 1 and 2

The Graph below shows the relationship between Tooth length and Dose

```
ggplot(aes(x=factor(dose), y=len), data=ToothGrowth) + geom_boxplot(aes(fill= factor(dose))) + ggtitle(
```



T-test for dose 0.5 mg:

```
t.test(len ~ supp, ToothGrowth[ToothGrowth$dose == .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 between group OJ and group VC 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
```

T-test for dose 1 mg:

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

```
##
##  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 between group OJ and group VC 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
```

T-test for dose 2 mg:

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

```
##
##  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 between group OJ and group VC 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
```

```
dose <- c(0.5, 1.0, 2.0)
p_value <- c(0.0064, 0.0010, 0.9639)
conf.int <-c("1.72, 8.78","2.80, 9.06","-3.80, 3.64")
decision <- c("Reject null", "Reject null", "Do not reject null")

knitr::kable(data.frame(dose, conf.int, p_value, decision), align = "cccc")
```

dose	conf.int	p_value	decision
0.5	1.72, 8.78	0.0064	Reject null
1.0	2.80, 9.06	0.0010	Reject null
2.0	-3.80, 3.64	0.9639	Do not reject null

As anticipated, the p-values for doses 0.5 and 1.0 are expected to be very low due to the substantial mean differences between them.

Consequently, for doses 0.5 and 1.0, since the p-values fall below 0.5, we can reject the null hypotheses asserting that there is no difference in tooth growth among the supplement types. However, for dose 2.0 mg/day, the null hypothesis can be retained as the p-value exceeds 0.5.

Conclusion

The fundamental assumption underlying the results is that the sample is a representative depiction of the population, and the variables are independent and identically distributed (IID) random variables.

Regarding the t-test, two key assumptions are taken into account:

1. The data is not paired, indicating independence.
2. The variances are unequal.

Given these considerations, upon reviewing the t-test results, it is observed that supplement type OC proves to be more effective than VC for doses below 1.0. However, at a dose of 2.0 mg/day, there is no discernible difference between the supplement types.