AMS 380.01: Quiz 1

Due on 02/25

Wei Zhu

Harris Temuri

Problem 1

Solution

(a) Please visualize the two groups using the Box plot in R.

```r
ggplot(data = data_1, aes(x = group, y = weight)) + geom_boxplot()
```

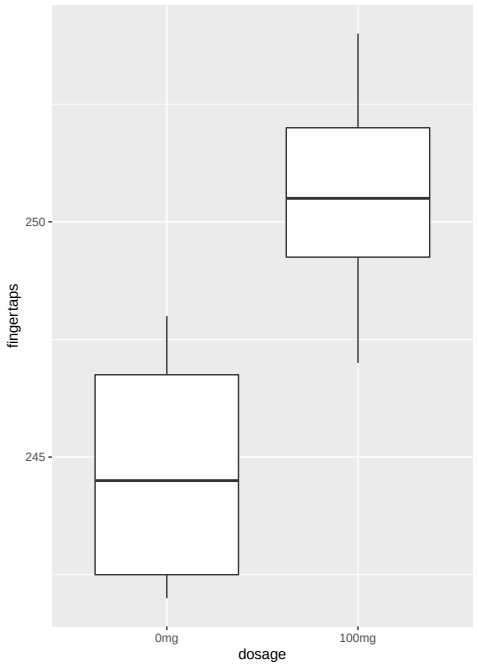

Figure 1: Box Plot of the Data_1

(b) Please test the normality and equal variance assumptions using R.

P-value of Shapiro test is greater than 0.05 for both 0mg and 100mg, so data is normally distributed.

```r
# Normality Test
# for 0mg
shapiro.test(data_1$fingertaps[data_1$dosage=="0mg"])
# for 100mg
shapiro.test(data_1$fingertaps[data_1$dosage=="100mg"])
```

Variance is equal (p-value above 0.05)

```
# Equal variance assumptions test
var.test(fingertaps ~ dosage, data=data_1)
```

(c) Please perform the correct test comparing whether the two means are equal or not using R. Please report the p-value of your test. Please also construct the 95% confidence interval for the mean difference using R.

```
# 1.3 2 sample t-test to test for equal means
t.test(data_1$fingertaps[data_1$dosage=="0mg"], data_1$fingertaps[data_1$dosage=="
    100mg"], var.equal = TRUE)
```

The p-value of the t-test is 2.402e-05 which is less than 0.05.

The 95% confidence interval is $[-7.825291 - 3.574709]$

(d) Please perform the correct test comparing whether the two means are equal or not by hand. Please make decision for you var.test(weight $\sim$ group, data = data_1) r hypothesis test at the significance level of $\alpha = 0.05$. Please also construct the 95% confidence interval for the mean difference by hand.
Mean difference is not 0 because p is significant

T-test:

$$t = \frac{m_A - m_B}{\sqrt{\frac{S^2}{n_A} + \frac{S^2}{n_B}}}$$

$$S^2 = \frac{\sum(x - m_A)^2 + \sum(s - m_B)^2}{n_A + n_B - 2}$$

$$S^2 = \frac{51.6 + 40.5}{18} = 5.116666667$$

$$t = \frac{250.5 - 244.8}{\sqrt{5.116666667/10 + 5.116666667/10}}$$

$$t = 5.634641574, p << 0.05$$

Confidence Interval:

$$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$250.5 - 244.8 \pm t_{0.05/2} \sqrt{5.116666667} \sqrt{\frac{1}{10} + \frac{1}{10}}$$

$$5.7 \pm (2.120)(2.262)(0.447213595)$$

$$5.7 \pm 2.144585962$$

$$[3.5554, 7.844585962]$$

Problem 2

Solution

(a) Please visualize the two groups using the Box plot and the Profile plot in R.

```r
# Box Plot
ggplot(data = data_2, aes(x = group, y = weight)) + geom_boxplot()
# Profile Plot
before <- subset(data_2, group=="before", weight, drop = TRUE)
after <- subset(data_2, group=="after", weight, drop = TRUE)
pd <- paired(before, after)
plot(pd, type="profile")
```

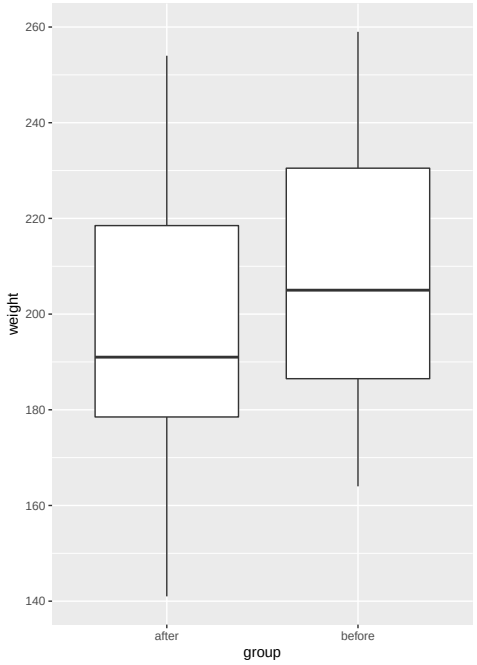

Figure 2: Box Plot of the Data_2

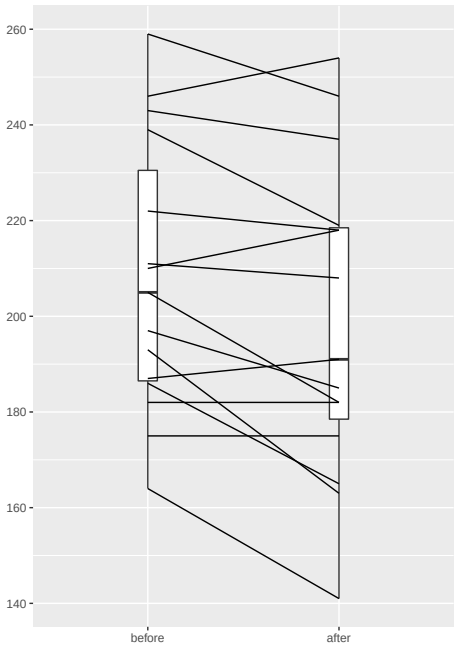

Figure 3: Profile Plot of the Data_2

(b) Please test the normality assumption using R.

P-value of Shapiro test is greater than 0.05 so data is normally distributed.

```
    # Normality Test
shapiro.test(data_2$weight[data_2$group=="after"] - data_2$weight[data_2$group=="before
    "])
```

(c) Please perform the correct test examining whether the mean difference is zero or not using R. Please report the p-value of your test. Please also construct the 95% confidence interval for the mean difference using R.

```
    # Paired t-test to test for equal means
    t.test(before, after, paired=TRUE)
```

The p-value of the paired t-test is 0.01299 which is less than 0.05.

The 95% confidence interval is $[2.21436, 15.78564]$

(d) Please perform the correct test examining whether the mean difference is zero or not by hand. Please make decision for your hypothesis test at the significance level of $\alpha = 0.05$. Please also construct the 95% confidence interval for the mean difference by hand.

Mean difference is not 0 because p is significant

Paired t-test:

$$t = \frac{m}{s/\sqrt{n}}$$

$$t = \frac{9}{12.25328/\sqrt{15}}$$

$$t = 2.844680448, p \approx 0.015$$

Confidence Interval:

$$\bar{x}_1 - \bar{x}_2 \pm t_{n-1,\alpha/2}(s_d/\sqrt{n})$$

$$9 \pm (2.145)(12.25328/\sqrt{15})$$

$$9 \pm 6.786351$$

$$[2.2136, 15.786351]$$

Problem 3

Solution

(a) Please perform the correct test examining whether the proportion of skiers catching cold are equal between the Vitamin C and the Placebo groups, using R. Please report the p-value of your test. Please also construct the 95% confidence interval for the mean difference in proportions using R.

```
prop.test(x = c(17,31), n = c(159,140))
```

The p-value of the proportional test is 0.01129 which is less than 0.05.
The 95% confidence interval is $[0.02390506, 0.20511561]$

(b) Please perform the correct test examining whether the proportion of the proportion of skiers catching cold are equal between the Vitamin C and the Placebo groups, by hand. Please make decision for your hypothesis test at the significance level of $\alpha = 0.05$. Please also construct the 95% confidence interval for the mean difference in proportions by hand.
Proportion is not equal because p is significant

z-test:

$$z = \frac{0.2214286 - 0.1069182}{\sqrt{pq/n_A + pq/n_B}}$$

$$z = \frac{0.1145286}{\sqrt{(0.160535117)(1-0.160535117)/140 + (0.160535117)(1-0.160535117)/159}}$$

$$z = \frac{0.1145286}{\sqrt{0.001809597}}$$

$$z = 2.69274153, p - value << 0.05$$

Confidence Interval:

$$p_A - p_B \pm z_{\alpha/2}(SE)$$

$$SE = \sqrt{\frac{p_A(s_d)}{n_A} + \frac{p_B(s_d)}{n_B}}$$

$$SE = \sqrt{\frac{0.2214286(0.7785714)}{140} + \frac{0.1069182(0.8930818)}{159}}$$

$$SE = \sqrt{0.001231414 + 0.000600545}$$

$$0.1145286 \pm 1.96(0.042801393)$$

$$[0.0306286, 0.198419331]$$