HUDM6026 Homework_07 In-class Activity

Chenguang Pan & Seng Lei

Mar 22, 2023

Part 1

Your group work will involve the acupuncture data set. I've uploaded a csv file in the Misc folder in the Files section.

Data for this group work come from a randomized experiment to study the efficacy of acupuncture for treating headaches. Results of the trial were published in the British Medical Journal in 2004. You may view the paper at the following link:http://www.bmj.com/content/328/7442/744.full The data set includes 301 cases, 140 control (no acupuncture) and 161 treated (acupuncture). Participants were randomly assigned to groups. Variable names and descriptions are as follows:age; age in years sex; male = 0, female = 1 migraine; diagnosis of migraines = 1, diagnosis of tension-type headaches = 0 chronicity; number of years of headache disorder at baseline acupuncturist; ID for acupuncture provider group; acupuncture treatment group = 1, control group = 0 pk1; headache severity rating at baseline pk5; headache severity rating 1 year later

Your task in group work today is to be done in three parts.

Part 1 is to run the standard two-sample t-test to test if acupuncture significantly decreased headache pain in study participants. Explore the assumptions of the t-test by examining the data through graphs.

MY SOLUTION

```
> # load the dataset
> df <- read.csv("acupuncture.csv")
> head(df)
   id age sex migraine chronicity acupuncturist group   pk1       pk5 remission
1 104  32   1        1         14            12     0 16.00 15.33333         0
2 108  56   1        1         40             9     0 16.50 23.25000         0
3 112  45   1        1         27             9     1  9.25  6.25000         1
4 113  45   1        1         30             9     1 42.50 51.25000         0
5 114  49   1        1         49             9     1 24.25 25.25000         0
6 126  47   1        1         42             5     0 21.00 15.25000         0
>
> # subset the treatment and control group
> group_treat <- df$pk5[which(df$group==1)]
> group_control <- df$pk5[which(df$group==0)]
>
> # draw the graph of treatment and control group's effect
> library(tidyverse)
> par(mfrow=c(1, 2))
> df %>% ggplot(aes(x = factor(group), y = pk5)) +
+   geom_boxplot()
> df %>% ggplot(aes(x = pk5),) +
+   geom_density(aes(color = factor(group)))
```
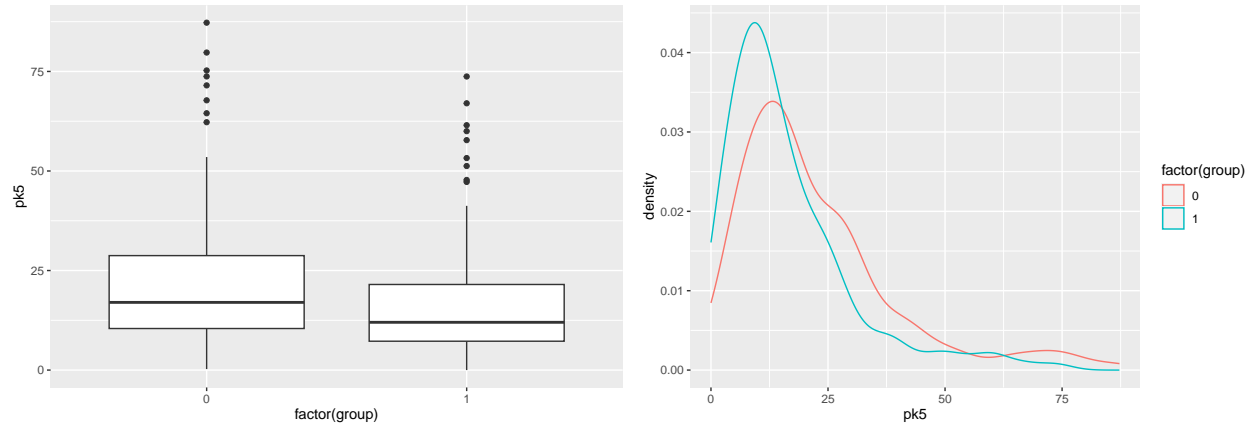
From the graphs we can see that the treatment group has lower average PK5 than control group. Next, I run the two-sample t-test and turn the argument `var.equal` to `FALSE`, although Levene's Test shows that there is no significant difference between the two groups' variance.

```
> # run the two-sample t-test
> t.test(group_treat, group_control,var.equal = FALSE)

    Welch Two Sample t-test

data:  group_treat and group_control
t = -3.3892, df = 266.6, p-value = 0.0008069
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -9.638399 -2.554924
sample estimates:
mean of x mean of y
 16.24679  22.34345
```

The Welch two sample t-test shows that there is statistically significant difference in two groups' average effect on PK5, $t(266.6) = -3.389$ and $p < .001$.

In addition, I checked the baseline balance on PK1 to ensure that we can use PK5 to ensure the ramdomization worked as intended.

```
> # install the package to run Cohen's d
> # install.packages("lsr")
> library(lsr)
> # extract the treatment and control group at baseline
> basegroup_treat <- df$pk1[which(df$group==1)]
> basegroup_control <- df$pk1[which(df$group==0)]
> # calculate the cohen's d to check the balance
> cohensD(basegroup_treat,basegroup_control)
[1] 0.1384973
```

The results show that the value of Cohen's D at baseline is .14, larger than .1 but fewer than .2. As a rule of thumb introduced by Maxwell et al.(2018), this result might cause some concern but not problematic. Further scrutiny is necessary. I also compare the ratio of group variance at baseline.

```
> # check the ratio of group variance for two groups at baseline
> ratio_of_variance <- var(basegroup_treat)/var(basegroup_control)
> ratio_of_variance
[1] 0.7079415
```

The ratio is smaller than .8, which may also be a sign of concern. As Prof. Keller's instruction in HUDM5123, we need to continue to check the covariates balance at baseline. But I ignored this issue and moved to the next task here.

Part 2

Part 2 is to identify a non-standard test statistic that your team suspects will also be sensitive to departures from the null hypothesis of no treatment effect. leveene's test of the variance ratio or correlation pk1-pk5 in treatment/ pk1-pk5 in control rank all the data in pk5

MY SOLUTION
First, I test the normality of two groups' distributions on PK5.

```
> # to test the normality of the two distribution
> shapiro.test(group_treat)

    Shapiro-Wilk normality test

data:  group_treat
W = 0.83345, p-value = 2.928e-12
> shapiro.test(group_control)

    Shapiro-Wilk normality test

data:  group_control
W = 0.84781, p-value = 1.017e-10
> # draw the qq plot
> mfrow=c(1,2)
> qqnorm(group_treat)
> qqline(group_treat)
> qqnorm(group_control)
> qqline(group_control)
```
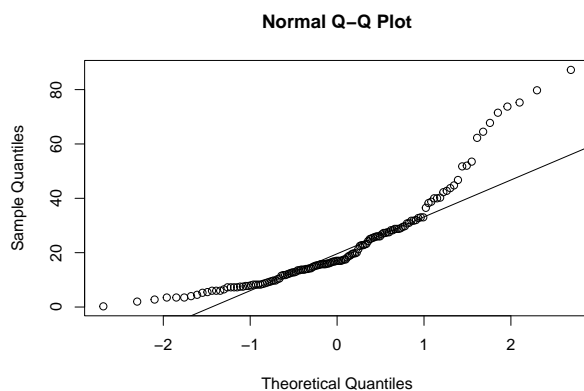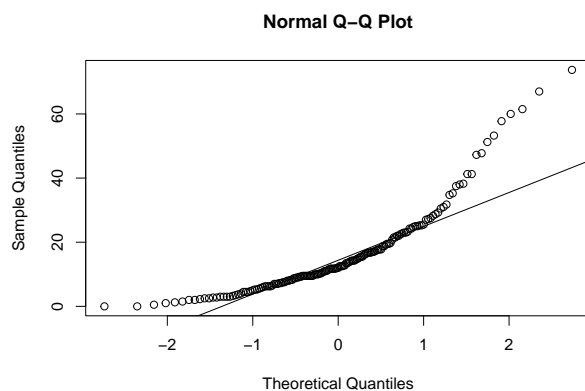

The Shapiro-Wilk Test and the Q-Q plot indicate that the both groups are not follow the normal distribution. From the original distribution graph made in the Part 1, one can see both distributions are positively skewed.

Therefore, running the regular two-sample T test in Part-1 may not a good choice since it compares the central tendency by using means and this statistic might be influenced by the extreme values.

We decided to choose the ranked-based test,i.e., the Wilcoxon rank-sum test, which is more robust to departures from normality since this test use median rather than mean to represent the central tendency and median is less sensitive to extreme values.

```
> obs_test_resulst <- wilcox.test(group_treat, group_control, alternative = "two.sided")
> obs_test_resulst

    Wilcoxon rank sum test with continuity correction

data:  group_treat and group_control
W = 8392, p-value = 0.000133
alternative hypothesis: true location shift is not equal to 0
```

The Wilcoxon rank-sum test's result presents that the two groups' medians are significantly different, $p < .001$. The two groups have different distributions.

Part 3

Part 3 is to use the non-standard test statistic in a permutation framework to determine the approximate significance level (p-value).

MY SOLUTION

```
> library(boot)
>
> # write a function to conduct the Wilcoxon rank-sum test
> median_diff <- function(dat, index) {
+    # fix the group assignments
+    treat_or_not <- dat$group
+    # permute the pk5 (combined data)
+    pk5 <- dat$pk5[index]
+    # extract the treat and control groups
+    treat <- pk5[which(treat_or_not==1)]
+    control <- pk5[which(treat_or_not==0)]
+    # get the difference in medians, and return the absolute value
+    test_out <- abs(median(treat)- median(control))
+    return(test_out)
+ }
>
>
> boot_out <- boot(data = df,
+                  statistic = median_diff,
+                  R = 10000, sim = "permutation")
> # use the observed statistic got from the part
> hist(boot_out$t, 50)
```

Histogram of boot_out$t
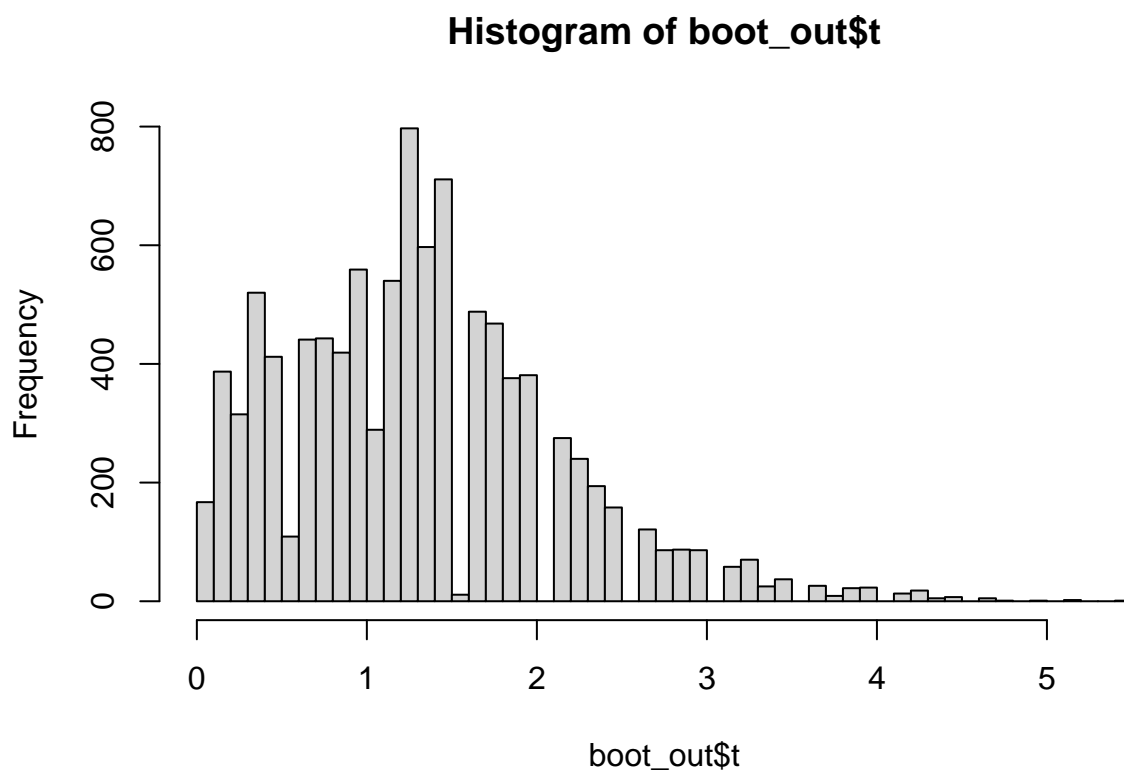

```
> p_value <- length(which(abs(boot_out$t) >= abs(boot_out$t0)))/ length(boot_out$t)
> p_value
[1] 4e-04
```

The permutation result indicates that the difference in medians is still statistically significant , $p=4 \times 10^{-4}$.