Homework 3

Yuanyou Yao

March 30, 2021

Problem 1

We first think the differences between two periods. Here we have:

$$-7, 3, -5, 4, -13, -6,$$

Correspondingly, the signs are:

$$-, +, -, +, -, -.$$

The p-value is

$$p = P\{\text{signs} \leq 2\} = 0.34$$

Therefore, we do not reject the null hypothesis.

Problem 2

We first think the rank of differences. Here we have:

$$-7, 3, -5, 4, -13, -6,$$

The ranks are:

$$5, 1, 3, 2, 6, 4$$

Correspondingly, the signs are:

$$-, +, -, +, -, -.$$

The total of positive ranks is $T_+ = 3$, and the total of negative ranks is $T_- = 18$.

p-value is $P\{T_+ \leq 3\} = P\{T_- \geq 18\} = 0.078$. Therefore we do not reject the null hypothesis.

Problem 3

Compare groups 20-29 vs 30-39. We firstly combine their data and rank the data

$$2, 8, 1, 6, 3.57, 3.5, 5$$

The total rank of group 20-29 is 17. So the p-value is

$$p = 2 \times P\{T \leq 16 = 0.69\}.$$

Therefore, we do not reject the null hypothesis.

We do the same calculation for 5 times more, and all of them suggest that we should not reject the null hypothesis.

Problem 4

We will rank the height in all age groups. We have:

Group 20-29: $2, 14, 1, 11$,

Group 30-39: $6, 12.5, 6, 10$,

Group 40-49: $15, 3, 4, 9$,

Group 50-59: $8, 6, 12.5$.

The total of ranks is $28, 34.5, 31, 26.5$ respectively.

Then we can calculate the statistic

$$H_{obs} = 0.40,$$

Since $H \approx \chi^2$, according to table in textbook, the p-value is

$$p = 0.95.$$

Thus the null hypothesis is retained.

Problem 5

As codes shown below, when degrees of freedom is 5, the null hypothesis is accepted. It is safe to say that t distribution can be approximated by the standard normal distribution.

```
x=rnorm(1000)
y=rt(1000,df = 4)
ks.test(x,y)

Two-sample Kolmogorov-Smirnov test

data:  x and y
D = 0.071, p-value = 0.01293
alternative hypothesis: two-sided

y=rt(1000,df = 5)
ks.test(x,y)

Two-sample Kolmogorov-Smirnov test

data:  x and y
D = 0.054, p-value = 0.1083
alternative hypothesis: two-sided
```