**Quantitative Methods**

**List of Exercises N. 6**

**Selected Exercises from McClave (2014) – Chapters 8**

* 1. **Comparing Two Population Means: Independent Sampling**

**Exercise 1. (22, SMILE). *Service without a Smile*. “Service with a smile” is a slogan that many businesses adhere to. However, there are some jobs (e.g., those of judges, law enforcement officers, pollsters) that require neutrality when dealing with the public. An organization will typically provide “display rules” to guide employees on what emotions they should use when interacting with the public. A *Journal of Applied Psychology* (Vol. 96, 2011) study compared the results of surveys conducted using two different types of display rules: positive (requiring a strong display of positive emotions) and neutral (maintaining neutral emotions at all times). In this designed experiment, 145 undergraduate students were randomly assigned to either a positive display rule condition (n1 = 78) or a neutral display rule condition (n2 = 67). Each participant was trained on how to conduct the survey using the display rules. As a manipulation check, the researchers asked each participant to rate, on a scale of 1 = “strongly agree” to 5 “strongly disagree” the statement: “This task requires me to be neutral in my expressions.”**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Positive Display Rule** | | | | | | | | | | | | | | | | | | | |
| **2** | **4** | **3** | **3** | **3** | **3** | **4** | **4** | **4** | **4** | **4** | **4** | **4** | **4** | **4** | **4** | **4** | **4** | **4** | **5** |
| **4** | **4** | **4** | **4** | **4** | **4** | **4** | **4** | **4** | **4** | **4** | **4** | **4** | **4** | **4** | **5** | **5** | **5** | **5** | **5** |
| **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** |
| **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** | **5** |  |  |
| **Neutral Display Rule** | | | | | | | | | | | | | | | | | | | |
| **3** | **3** | **2** | **1** | **2** | **1** | **1** | **1** | **2** | **2** | **1** | **2** | **2** | **2** | **3** | **2** | **2** | **1** | **2** |  |
| **2** | **2** | **2** | **2** | **1** | **2** | **2** | **2** | **2** | **2** | **2** | **1** | **2** | **2** | **2** | **2** | **2** | **2** | **2** |  |
| **3** | **2** | **1** | **2** | **2** | **2** | **1** | **2** | **1** | **2** | **2** | **3** | **2** | **2** | **2** | **2** | **2** | **2** | **2** |  |
| **2** | **2** | **2** | **2** | **1** | **2** | **2** | **2** | **2** | **2** |  |  |  |  |  |  |  |  |  |  |

First, we need to do as followed:

1. Clean your screen
2. Run Library(mosaic)
3. Import data
4. Create variables

For explanation how to do this look at List 1, Exercise 1.

Clean the R Script: rm(list=ls())

Library: library(mosaic)

Import data: library(readxl)

In this exercise, we have two datasets.

L6E1A <- read\_excel("~/SDC/Quantitative Research Methods Course/Exercises/Dataset/List6/L6E1.xlsx")

View(L6E1)

attach(L6E1)

Create variables: POSITIVE <- POSITIVE

NEUTRAL <- NEUTRAL

1. **If the manipulation of the participants was successful, which group should have the larger mean response? Explain.**

If the manipulation was successful, then the positive group (requiring a strong display of positive emotions) should have the higher mean response. The members of this group should disagree with the statement presented, resulting in higher responses.

1. **The data for the study (simulated based on information provided in the journal article) are listed in the table below. Access the data and run an analysis to determine if the manipulation was successful. Conduct a test of hypothesis using α = 0.05.**

Let µ1 = mean response for the positive group and µ2=mean response for the neutral group. To determine if the manipulation was successful, we test:

Important!! We only check for bigger means, as this is how the experiment manipulated the participants. Therefore, this is a one-tail test.

Remove NA’s for each of the variables:

na.omit(VARIABLE)

First, we need to calculate some descriptive statistics:

> mean(POSITIVE)

[1] 4.487179

> mean(NEUTRAL1)

[1] 1.895522

> var(POSITIVE)

[1] 0.4348984

> var(NEUTRAL1)

[1] 0.2464948

> length(POSITIVE)

[1] 78

> length(NEUTRAL1)

[1] 67

The test statistic is:

In R:

2.591657/sqrt(0.009254648)= 26.94

The rejection region requires alpha = 0.05 in the upper tail of the z-distribution, as this is a one-tail test. From the z-table, we find z=1.645. The rejection region is thus z>1.645.

Since our test statistic falls in the rejection region, we reject the null hypothesis. There is sufficient evidence to indicate that the mean response for the positive group is bigger than the mean response for the neutral group.

1. **What assumptions, if any, are required for the inference from the test to be valid?**

We need to assume that random and independent samples were selected from each of the populations.

We assume the samples have normal distributions because of the central limit theorem.

**Exercise 2. (24, GASTRB). *Cooling method for gas turbines*. The *Journal of Engineering for Gas Turbines and Power* (Jan. 2005) dis a study of gas turbines augmented with high-pressure inlet fogging. The researchers classified gas turbines into 3 categories: traditional, advanced, and aeroderivative. Summary statistics on heat rate (kilojoules per kilowatt per hour) for each of the 3 types of gas turbines are shown in the table below.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Engine** | **N** | **Mean** | **St Dev** | **Minimum** | **Maximum** |
| **Heatrate** | **Advanced** | **21** | **9764** | **639** | **9105** | **11588** |
|  | **Aeroderiv** | **7** | **12312** | **2652** | **8714** | **16243** |
|  | **Traditional** | **39** | **11544** | **1279** | **10086** | **14796** |

1. **Is there sufficient evidence of a difference between the mean heat rates of traditional augmented gas for turbines and aeroderivative augmented gas turbines? Test using an α = 0.05.**

Let µ1= the mean heat rates of traditional augmented gas turbines and µ2= the mean heat rates of aeroderivative augmented gas turbines.

We use the Students t-test, as the true variance of the populations from the population which the samples are extracted from is unknown. Therefore, we have to do some extra calculations compared to the exercises before, which was a large sample test.

Some preliminary calculations are:

To determine if there is a difference in the mean heat rates for traditional augmented gas turbines and the mean heat rates of aeroderivative augmented gas turbines, we test:

The test statistic is:

In R:

((11544-12312)-0)/(sqrt(2371831.409\*((1/39)+(1/7))))

Result: t = -1.214847

The rejection region requires α/2=0.05/2=0.025 in each tail of the t-distribution with:

df=n1+n2-2

df=39+7-2=44

From the t table, we can see that t0.0252.021. The rejection region is:

t<-2.021 or t>2.021.

Since the observed value of the test statistic does not fall in the rejection region (t=-1.20 differ from -2.021), H0 is not rejected. There is insufficient evidence to indicate that there is a difference in the mean heat rates for traditional augmented gas turbines and the mean heat rates of aeroderivative augmented gas turbines at α = 0.05.

1. **Is there sufficient evidence of a difference between the mean heat rates of advanced augmented gas turbines and aeroderivative augmented gas turbines? Test using an α = 0.05.**

Let µ3= the mean heat rates of advanced augmented gas turbines and µ2= the mean heat rates of aeroderivative augmented gas turbines.

Some preliminary calculations are:

To determine if there is a difference in the mean heat rates for traditional augmented gas turbines and the mean heat rates of aeroderivative augmented gas turbines, we test:

The test statistic is:

In R:

((9764-12312)-0)/(sqrt(1937117.077\*((1/21)+(1/7))))

t=-4.194702

The rejection region requires α/2=0.05/2=0.025 in each tail of the t-distribution with:

df=n1+n2-2

df=21+7-2=46

From the t table, we can see that t0.0252.056. The rejection region is:

t<-2.056 or t>2.056.

Since the observed value of the test statistic does not fall in the rejection region (t=-4.19 < -2.056), H0 is rejected. There is sufficient evidence to indicate that there is a difference in the mean heat rates for advanced augmented gas turbines and the mean heat rates of aeroderivative augmented gas turbines at α = 0.05.

**8.3 Comparing Two Population means: Paired Difference sampling**

**Exercise 3. (37, NAPS). *Taking ”power naps” during work breaks*. Lack of sleep costs companies about 18 billion dollars a year in lost productivity according to the National Sleep Foundation. Companies are waking up to the problem, however. Some even have quiet rooms available for study or sleep. “Power naps” are in vogue (*Athens Daily News*, Jan. 9, 2000). A major airline recently began encouraging reservation agents to nap during their breaks. The accompanying table lists the number of complaints received about each of a sample of 10 reservation agents during the 6 months before naps were encouraged and during the 6 months after the policy change.**

|  |  |  |
| --- | --- | --- |
| **Operator** | **Before Policy** | **After Policy** |
| **1** | **10** | **5** |
| **2** | **3** | **0** |
| **3** | **16** | **7** |
| **4** | **11** | **4** |
| **5** | **8** | **6** |
| **6** | **2** | **4** |
| **7** | **1** | **2** |
| **8** | **14** | **3** |
| **9** | **5** | **5** |
| **10** | **6** | **1** |

First, we need to do as followed:

1. Clean your screen
2. Run Library(mosaic)
3. Import data
4. Create variables

For explanation how to do this look at List 1, Exercise 1.

Clean the R Script: rm(list=ls())

Library: library(mosaic)

Import data: library(readxl)

L6E3 <- read\_excel("~/SDC/Quantitative Research Methods Course/Exercises/Dataset/List6/L6E3.xlsx")

View(L6E3)

attach(L6E3)

Create variables: Operator <- Operator

Before <- Before

After <- After

1. **Do the data present sufficient evidence to conclude that the new napping policy reduced the mean number of customers complaints about reservations agents? Test using α = 0.05.**

In this exercise, we have two ways of calculating this.

To determine if the new napping policy reduced the mean number of customer complaints, we test:

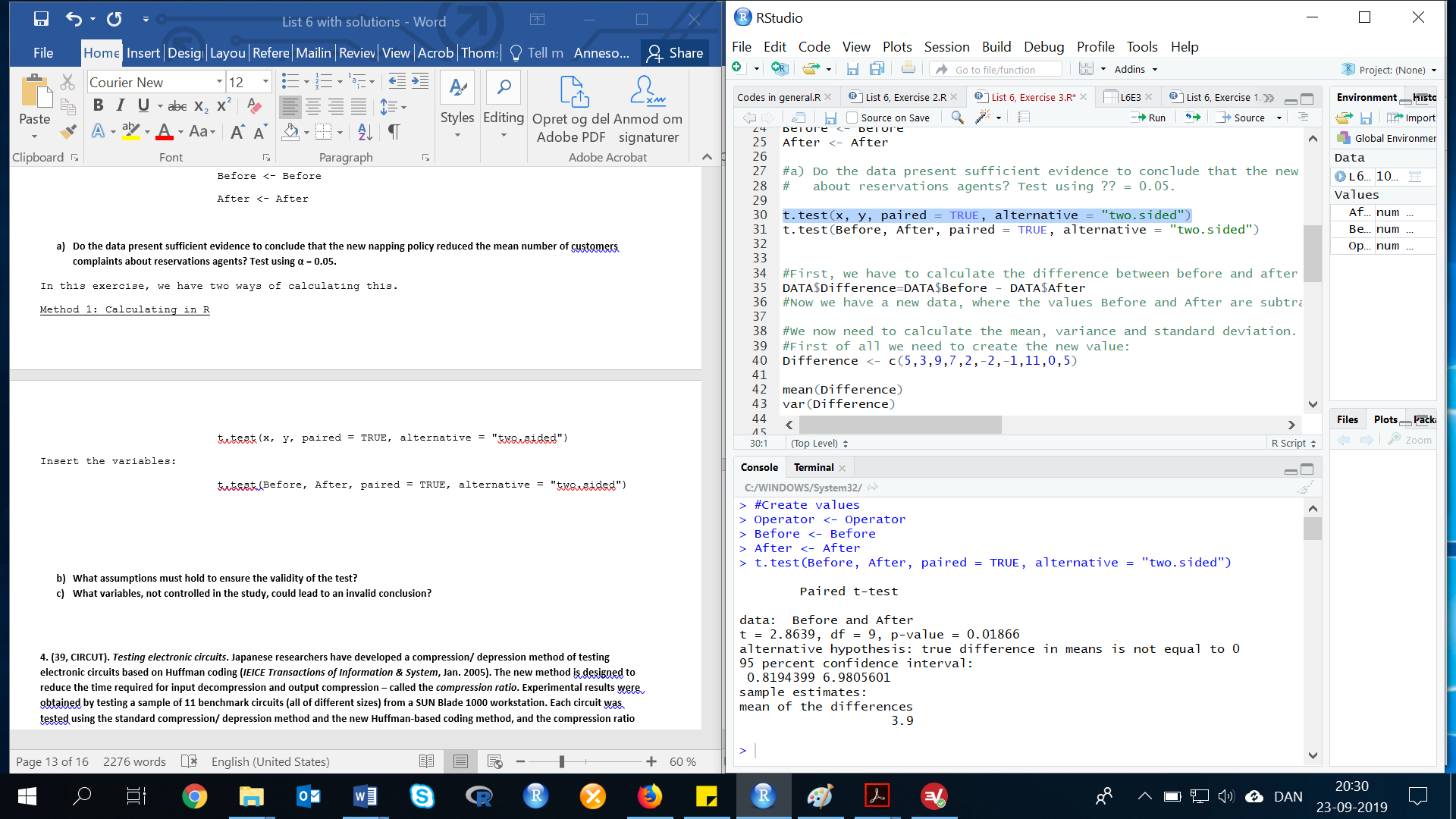
Method 1: Calculating with R

t.test(x, y, paired = TRUE, alternative = "two.sided")

Insert the variables:

t.test(Before, After, paired = TRUE, alternative = "two.sided")

The R Console give us the following result:



Result: t=2.8639.

Method 2: Manual calculating

First, we have to calculate the difference between before and after the new policy was introduced, by subtracting the value for "After" from "Before".

DATA$Before - DATA$After

DATA$Difference <- DATA$Before - DATA$After

Now we have a new data, where the values Before and After are subtracted. We can now calculate the mean, variance and standard deviation.

mean(Difference)

var(Difference)

sd(Difference)

Result:

Mean = 3.9

Variance = 18.54444

Standard deviation = 4.306326

The test statistics is:

t=(3.9-0)/(4.3063/(sqrt(10)))

Result: t=2.863916

The rejection region requires alpha=0.05 in the upper tail of the t-distribution with df=n(d)-1:

df=10-1

df=9

Both methods give the same result t=2.863913. From the t-distribution table we can see that, t0.05=1.833. The rejection region is: t>1.833. Since the observed value of the test statistic falls in the rejection region (t=2.864>1.833), H0 is rejected. There is sufficient evidence to indicate the new napping policy reduced the mean number of customer complaints at alpha=0.05.

1. **What assumptions must hold to ensure the validity of the test?**

In order for the above test to be valid, we must assume that:

1. The population of differences is normal.
2. The differences are randomly selected.
3. **What variables, not controlled in the study, could lead to an invalid conclusion?**

Variables that were not controlled that could lead to an invalid conclusion include time of day agents worked, day of the week agents worked, and how much sleep the agents got before working, among others.

**Exercise 4. (39, CIRCUT). *Testing electronic circuits*. Japanese researchers have developed a compression/ depression method of testing electronic circuits based on Huffman coding (*IEICE Transactions of Information & System*, Jan. 2005). The new method is designed to reduce the time required for input decompression and output compression – called the *compression ratio*. Experimental results were obtained by testing a sample of 11 benchmark circuits (all of different sizes) from a SUN Blade 1000 workstation. Each circuit was tested using the standard compression/ depression method and the new Huffman-based coding method, and the compression ratio was recorded. The data are given in the table below. Compare the two methods with a 95% confidence interval. Which method has the smaller mean compression rate?**

|  |  |  |
| --- | --- | --- |
| **Circuit** | **Standard Method** | **Huffman-coding Method** |
| **1** | **.80** | **.78** |
| **2** | **.80** | **.80** |
| **3** | **.83** | **.86** |
| **4** | **.53** | **.53** |
| **5** | **.50** | **.51** |
| **6** | **.96** | **.68** |
| **7** | **.99** | **.82** |
| **8** | **.98** | **.72** |
| **9** | **.81** | **.45** |
| **10** | **.95** | **.79** |
| **11** | **.99** | **.77** |

First, we need to do as followed:

1. Clean your screen
2. Run Library(mosaic)
3. Import data
4. Create variables

For explanation how to do this look at List 1, Exercise 1.

Clean the R Script: rm(list=ls())

Library: library(mosaic)

Import data: library(readxl)

L6E4 <- read\_excel("~/SDC/Quantitative Research Methods Course/Exercises/Dataset/List6/L6E4.xlsx")

View(L6E4)

attach(L6E4)

Create variables: Standard <- STANDARD

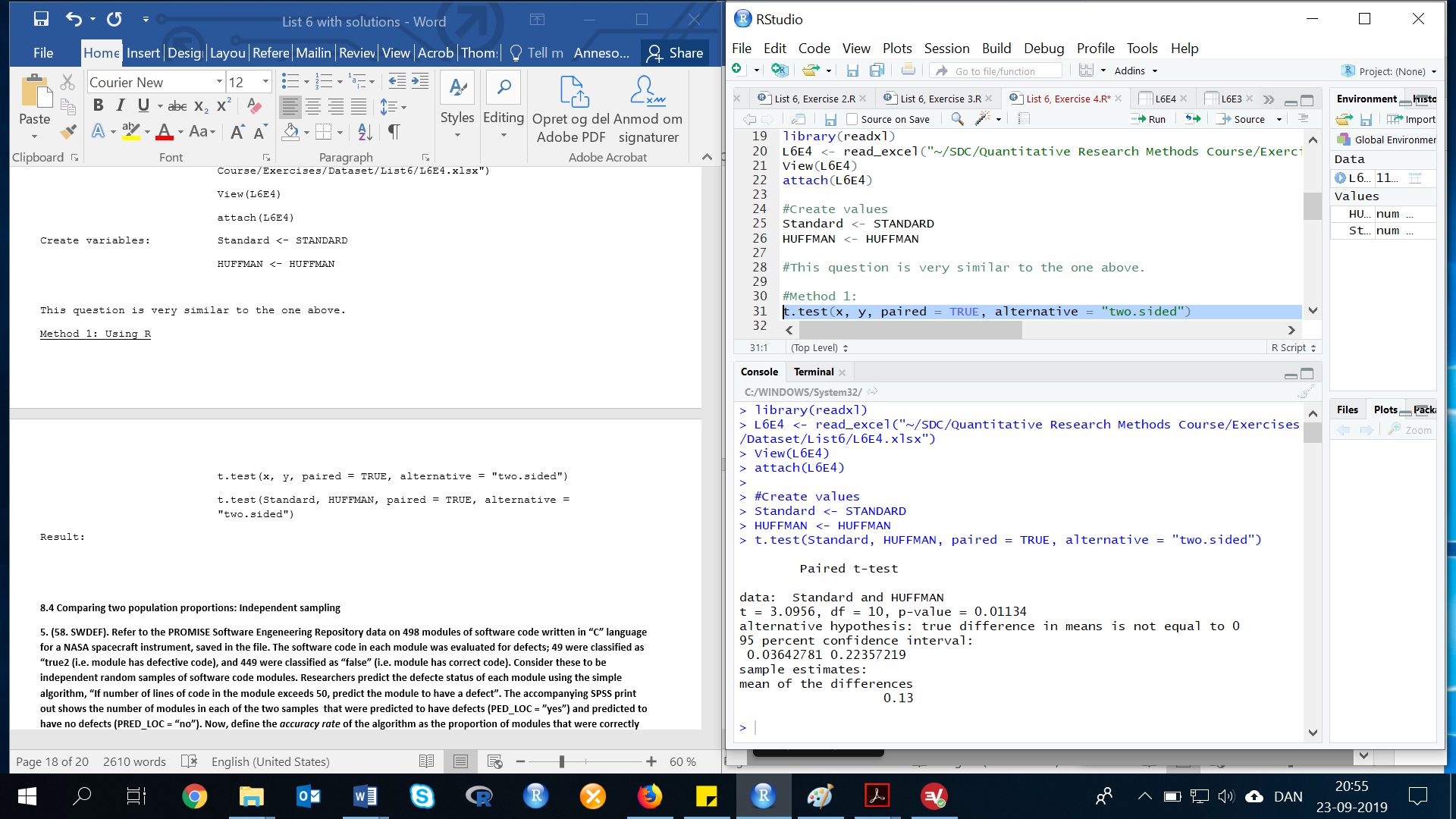
HUFFMAN <- HUFFMAN

This question is very similar to the exercise above.

Method 1: Using R

t.test(x, y, paired = TRUE, alternative = "two.sided")

t.test(Standard, HUFFMAN, paired = TRUE, alternative = "two.sided")

Result:

95% cofidence interval: (0.03643363,0.2235664)

We are 95% confident that the true difference in mean compression ratio between the standard method and the Huffman-based coding method is between 0.036 and 0.224. Since 0 is not contained in the interval, we can conclude there is a difference in mean compression ratios between the two methods. Since the values of the confidence interval are positive, we can conclude that the mean compression ratio for the Huffman-based method is smaller than the standard method.

Method 2: Manual calculation

we need to calculate the differences in compression rates and then the relevant descriptive statistics for the sample.

STANDARD - HUFFMAN

Difference <- STANDARD - HUFFMAN

Now we have a new data, where the values Standard and Huffman are subtracted. We will use this dataset to to calculate the mean, variance and standard deviation, but first of all we need to create the new value:

mean(Difference)

var(Difference)

sd(Difference)

Results:

mean = 0.13

variance = 0.0194

standard deviation = 0.1392839

For confidence coefficient 0.95, alpha=.05 and alpha/2=.05/2=.025. From the t distribution table, we can find t(.025), but we need to now the degress of freedom first, df=n-1.

df=11-1

df=10.

With df=10, can we see at the t-distribution table, that t(.025)=2.228. The 95% confidence interval is:

0.13+(2.228\*((0.1392839)/(sqrt(11))))

0.13-(2.228\*((0.1392839)/(sqrt(11))))

Result: (0.03643363, 0.2235664)

We are 95% confident that the true difference in mean compression ratio between the standard method and the Huffman-based coding method is between 0.036 and 0.224. Since 0 is not contained in the interval, we can conclude there is a difference in mean compression ratios between the two methods. Since the values of the confidence interval are positive, we can conclude that the mean compression ratio for the Huffman-based method is smaller than the standard method.

**8.4 Comparing two population proportions: Independent sampling**

**Exercise 5. (58. SWDEF). Refer to the PROMISE Software Engeneering Repository data on 498 modules of software code written in “C” language for a NASA spacecraft instrument, saved in the file. The software code in each module was evaluated for defects; 49 were classified as “true2 (i.e. module has defective code), and 449 were classified as “false” (i.e. module has correct code). Consider these to be independent random samples of software code modules. Researchers predict the defecte status of each module using the simple algorithm, “If number of lines of code in the module exceeds 50, predict the module to have a defect”. The accompanying SPSS print out shows the number of modules in each of the two samples that were predicted to have defects (PED\_LOC = ”yes”) and predicted to have no defects (PRED\_LOC = “no”). Now, define the *accuracy rate* of the algorithm as the proportion of modules that were correctly predicted. Compare the accuracy rate of the algorithm when applied to modules with defective code to the accuracy rate of the algorithm when applied to modules with correct code. Use 99% confidence interval.**

**DEFECT \* PRED\_LOC Crosstabulation**

**Count**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **PRED\_LOC** | | **Total** |
| **no** | **yes** |
| **DEFECT false**  **true**  **Total** | **400**  **29**  **429** | **49**  **20**  **69** | **449**  **49**  **498** |

Let p1=accuracy rate for modules with correct code and p2=accuracy rate for modules with defective code. Some preliminary calculations are:

Result: p1 = 0.8908686, p2 = 0.4081633

For confidence coefficient 0.99, α=0.01 and α/2=0.01/2=0.005. From the z table, we can see that . The 99% confidence interval is:

The 99% confidence interval is: (0.298, 0.668)

We are 99% confident that the difference in accuracy rates between modules with correct code and modules with defective code is between 0.298 and 0.668.

**8.5 Determining the required sample size**

**Exercise 6. (70). *Users of home shopping services*. All cable companies carry at least one home shopping channel. Who uses these home shopping services? Are the shoppers primarily men or women? Suppose you want to estimate the difference in the proportions of men and women the difference in the proportions of men and women who say they have used or expect to use televised home shopping using an 80% confidence interval of width .06 or less.**

**a) Approcimatelly how many people should be included in your samples?**

For confidence coefficient 0.80, α=0.20 and α/2=0.20/2=0.10. From the z table, we can see that z0.10=1.28. Since we have no prior information about the proportions, we use p1=p2=0.50 to get a conservative estimate. Look in the book at page 475, if you need furhter explanation. For a width of 0.06, the margin error is 0.03.

In R: ((1.28^2)\*((0.5\*(1-0.5))+(0.5\*(1-0.5))))/(0.03^2)

Result: n1=n2 = 910.2222

**b) Suppose you want to obtain individual estimates for the two proportions of interest. Will the sample size found in part a be large enough to provide estimates of each proportion correct to within .02 with probability equal to .90? Justify your response.**

For confidence coefficient 0.90, α=0.10 and α/2=0.10/2=0.05. From the z table, we can see that z0.05=1.645. Using the formula for the sample size needed to estimate a proportion. Look in the book at page 475, if you need furhter explanation.

In R: ((1.645^2)\*(0.5\*(1-0.5)))/(0.02^2)

n=1691.266

No, the sample size from part a is not large enough.

**8.6 Comparing two population variances: Independent sampling**

**Exercise 7. (88, HCOUGH). *Is honey a cough remedy?* Refer to the Archives of pediatrics and adolescent medicine (Dec. 2007) study of honey as a children cough remedy. Exercise 23. The data (cough improvement scores) for the 33 children in the DM dosage group and the 35 children in the honey dosage group are reproduced in the table below. In exercise 23, you used a comparison of two means to determine whether “honey may be a preferable treatment for the cough and sleep difficulty associated with childhood upper respiratory tract infection”. The researchers also want to know if the variability in coughing improvement scores differ for the two groups. Conduct the appropriate analysis, using α = .10.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Honey dosage:** | **12** | **11** | **15** | **11** | **10** | **13** | **10** | **4** | **15** | **16** | **9** | **14** | **10** | **6** | **10** | **8** | **11** | **12** | **12** | **8** |
| **9** | **11** | **15** | **10** | **15** | **9** | **13** | **8** | **12** | **10** | **8** | **9** | **5** | **12** | **12** |  |  |  |  |  |
| **DM dosage:** | **4** | **6** | **9** | **4** | **7** | **7** | **7** | **9** | **12** | **10** | **11** | **6** | **3** | **4** | **9** | **12** | **7** | **6** | **8** | **12** |
| **12** | **4** | **12** | **13** | **7** | **10** | **13** | **9** | **4** | **4** | **10** | **15** | **9** |  |  |  |  |  |  |  |

First, we need to do as followed:

1. Clean your screen
2. Run Library(mosaic)
3. Import data
4. Create variables

For explanation how to do this look at List 1, Exercise 1.

Clean the R Script: rm(list=ls())

Library: library(mosaic)

Create the data and variables in R:

HONEY <- c(12, 11, 15, 11, 10, 13, 10, 4, 15, 16, 9, 14, 10, 6, 10, 8, 11, 12, 12, 8, 9, 11, 15, 10, 15, 9, 13, 8, 12, 10, 8, 9, 5, 12, 12)

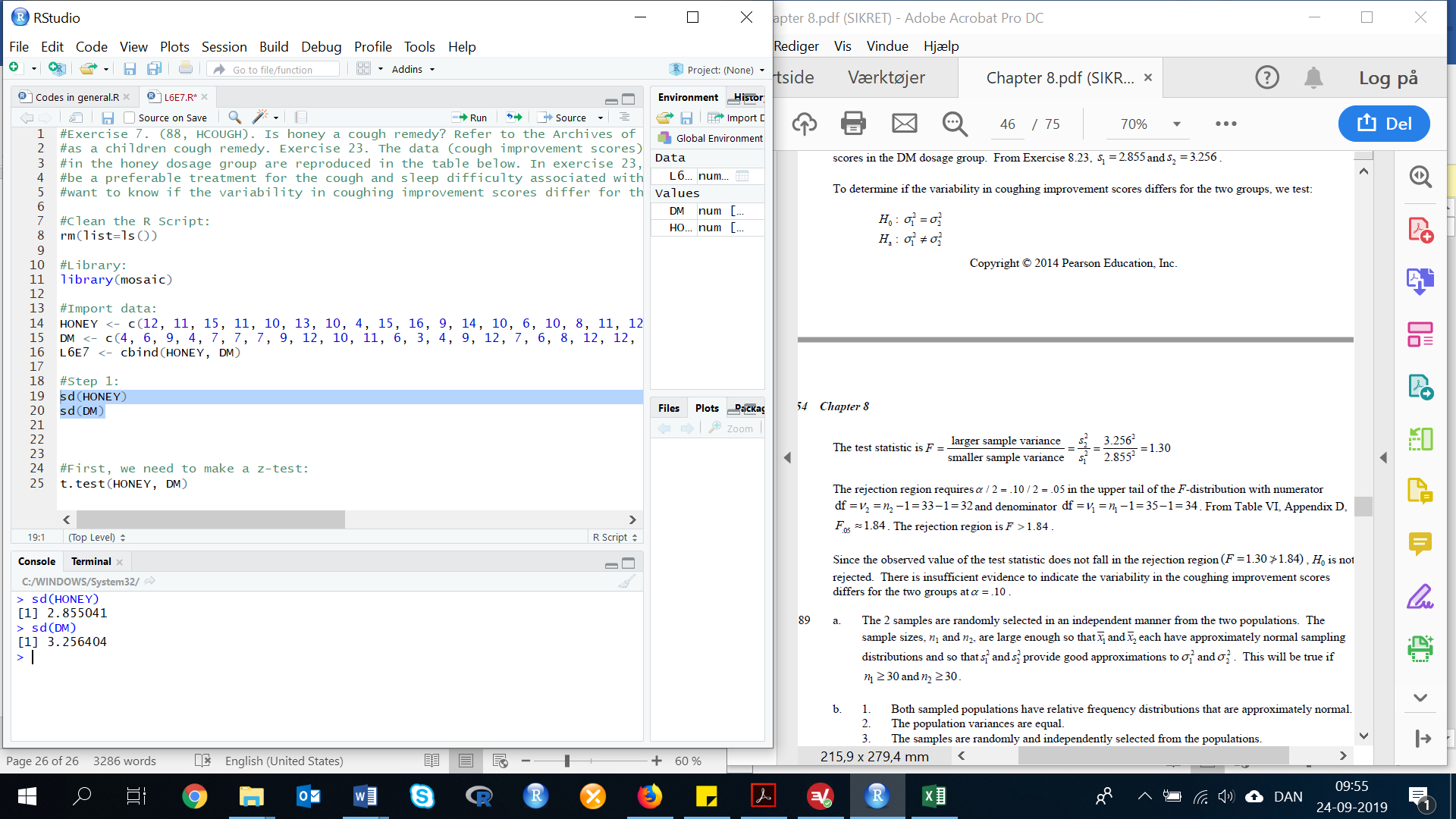
DM <- c(4, 6, 9, 4, 7, 7, 7, 9, 12, 10, 11, 6, 3, 4, 9, 12, 7, 6, 8, 12, 12, 4, 12, 13, 7, 10, 13, 9, 4, 4, 10, 15, 9)

L6E7 <- cbind(HONEY, DM)

Let =variance of improvement scores in the honey dosage group and =variance of improvement scores in the DM dosage group. To solve this exercise, we need to know the standard deviation of both Honey and DM.

Calculation of the standard deviation:

sd(HONEY)

sd(DM)

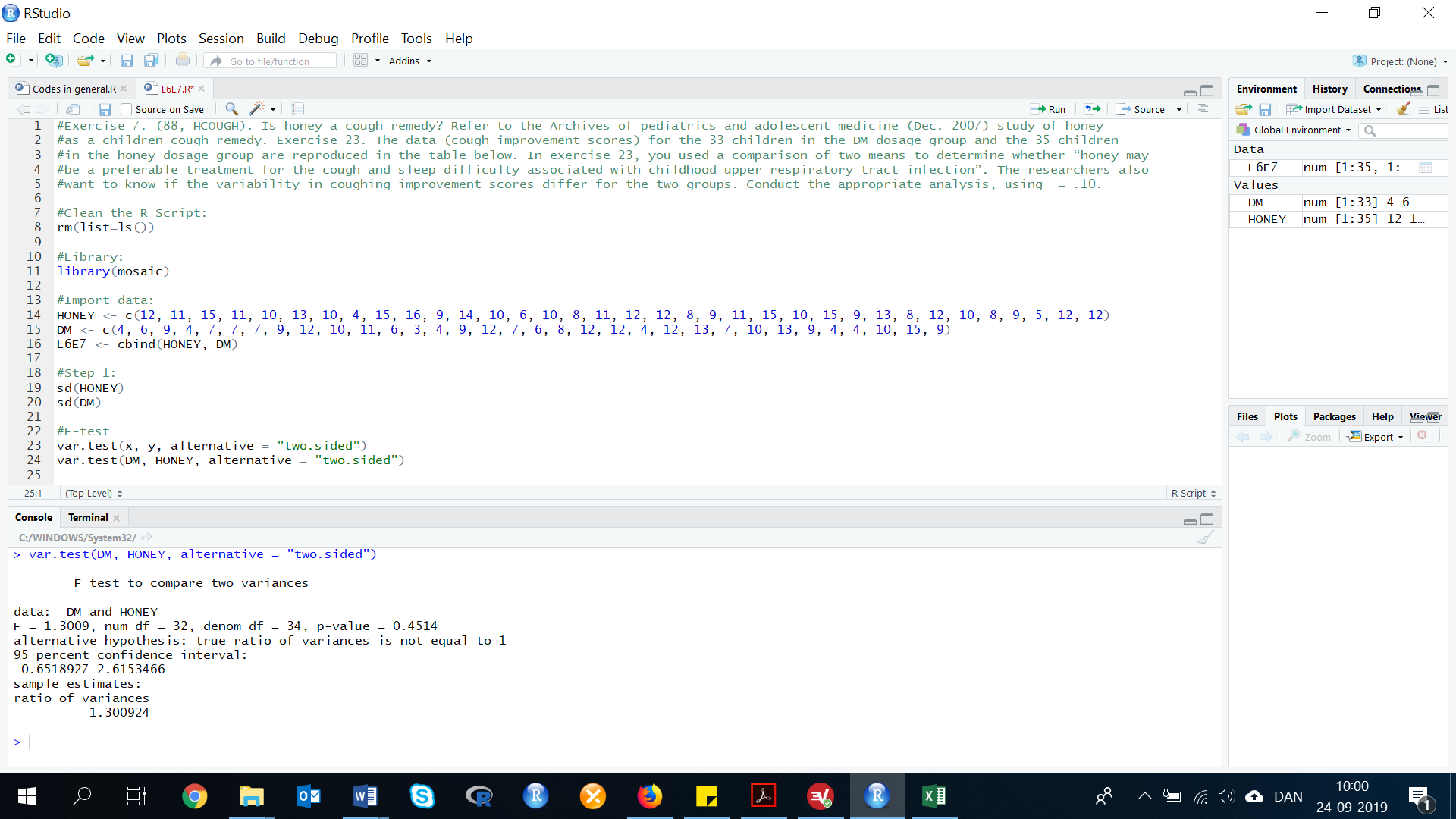
To determine if the variability in coughing improvement scores differs for the two groups, we test:

There is two ways to solve this exercise:

Method 1: Calculating with R

The test statistic is: var.test(x, y, alternative = "two.sided")

var.test(DM, HONEY, alternative = "two.sided")

The R Console gives us the following result:

Therefore, F = 1.300924.

The rejection region requires α/2=0.10/2=0.05 in the upper tail of the F-distribution with numerator df=v2-n2-1=33-1=32 and denominator df=v1-n1-1=35-1=34. From the F distribution table, we can see that F0.05=1.84. The rejection region is F>1.84.

Since the observed value of the test statistic does not fall in the rejection region (F=1.30 differs from 1.84), H0 is not rejected. There is insufficient evidence to indicate the variability in the coughing improvement scores differs for the two groups at α=0.10.

Method 2: Manual calculation

The test statistic is:

F = 1.30.

The rejection region requires α/2=0.10/2=0.05 in the upper tail of the F-distribution with numerator df=v2-n2-1=33-1=32 and denominator df=v1-n1-1=35-1=34. From the F distribution table, we can see that F0.05=1.84. The rejection region is F>1.84.

Since the observed value of the test statistic does not fall in the rejection region (F=1.30 differs from 1.84), H0 is not rejected. There is insufficient evidence to indicate the variability in the coughing improvement scores differs for the two groups at α=0.10.