Johnny Chen

R Assignment 8

**Make sure you do the following:**

1. **Put your name in the left hand corner of your document.**
2. **Put answers in document.**
3. **Use complete and coherent sentences to answer the questions.**
4. **Save file with your name R 7 (e.g. Chimiak\_S\_R\_8)**
5. **Upload completed R Assignment on Canvas.**
6. The data set can be found on Blackboard in R\_8\_Data (Column 1, titled Heat Conductivity).

How well materials conduct heat matters when designing houses, for example. Conductivity is measured in terms of watts of heat [**power**](JavaScript:top.Define('power')) transmitted per square meter of surface per degree Celsius of temperature difference on the two sides of the material. In these units, glass has a conductivity of about 1. The National Institute of Standards and Technology (NIST) provides data on properties of materials. Here is a simple random sample of 11 NIST measurements of the heat conductivity of a particular type of glass:

1.11 1.07 1.11 1.07 1.12 1.08 1.08 1.15 1.18 1.18 1.12

(Remember this data is in Column 1)

* 1. **(1)**Make a probability plot (0.25pts) and a boxplot (0.25pts) and see if there are any major deviations from Normality. Is it reasonable to use the t procedures (0.25pts)? Why or why not (0.25 pt)? (Copy and paste the graphs that result with your answer. Make sure you have meaningful titles for your graph.)

***If the answer is no, do not do parts b, c, and d.***

**Help on Making probability plots and boxplots**

**Generic R Code and an example of the code for Problem 1, you can use this to guide you with the other problems, for a Probability Plot:**

qqnorm(DATA\_SET\_NAME$Variable\_Name,main=”Input your title”)

#Code for problem 1 part abelow

qqnorm(R\_8\_Data$Heat\_Conductivity,main="Heat Conductivity")

**R Code for a Boxplot:**

Boxplot(DATA\_SET\_NAME$Variable\_Name,main=”Input your title”,xlab=”label for x-axis”,horizontal=TRUE)

#Code for problem 1 part a below

boxplot(R\_8\_Data$Heat\_Conductivity, main="Distribution of Length of Conductivity for Sample of 11 of a Particular Class",xlab="Conductivity",horizontal=TRUE)

***Remember, If the answer is no, do not do parts b, c, and d. Help on doing confidence intervals and hypothesis tests are shown after questions b and c***

***A graph with numbers and lines

AI-generated content may be incorrect.*** ***A screenshot of a computer

AI-generated content may be incorrect.***

Yes, it seems reasonable to use t procedures for this problem. Our qq-plot appears approximately linear, so it’s fair to assume that our distribution is approximately normal. Furthermore, our boxplot shows there are no outliers present in the dataset. Thus we conclude that it is reasonable to use t procedures.

* 1. **(1 pts, if needed to be done or -0.5 if you do and it didn’t need to be done)** **If** your conclusion in part (a) is **yes**, give a 92% confidence interval for the mean conductivity, round answer to 4 decimal places. (Copy and paste results. Make sure your interpretation of the 92% confidence interval is in a complete sentence in the context of the problem.)

A white background with black text

AI-generated content may be incorrect.

We are 92% confident that the mean conductivity of the glass lies between 1.0918 and 1.1391.

* 1. (**3 pts, if needed to be done or -1.5 if you do and it didn’t need to be done**) **If** your conclusion in part (a) is **yes**, do the data give convincing evidence that the mean heat conductivity of this particular type of glass is less than 1.15 at a 0.01 significance level?
     1. (0.5) State the null and alternative hypotheses.

Null Hypothesis: The mean heat conductivity of the glass is equal to 1.15.

Alternative Hypothesis: The mean heat conductivity of the glass is less than 1.15.

* + 1. (0.5) State the significance level for this problem.

The significance level is 0.01.

* + 1. (0.5) State the test statistic to 4 decimal places. Use R to compute the test statistic and *P*-value. (Copy and paste the results. Also, use this result to answer part (iv)).

A white background with black text

AI-generated content may be incorrect.

The test statistic is -2.8403.

* + 1. (0.5) State the *P*-value to five decimal places.

The p-value is 0.00877.

* + 1. (0.5)State whether you reject or do not reject the null hypothesis.

Because 0.00877 < 0.01, we reject the null hypothesis.

* + 1. (0.5)State your conclusion in context of the problem.

We conclude that the mean heat conductivity of the glass is less than 1.15 at a 0.01 significance level.

Confidence Intervals and Hypothesis Testing

#Code to get Confidence Interval

t.test(DATA\_Set\_name$Variable\_Name, conf.level = confidence level in decimal form)

#Code for problem 1 part b below

t.test(R\_8\_Data$Heat\_Conductivity, conf.level = 0.92)

#Code to perform a hypothesis test

t.test(DATA\_Set\_name$Variable\_Name, mu= null\_hypothesis\_value, alternative = "two.sided" or "less" or "greater")

#Code for problem 1 part c below

t.test(R\_8\_Data$Heat\_Conductivity,mu=1.15,alternative= "less")

1. The data set is in (Columns 3, 4, and 5 labeled Before, After, and Difference\_2).

Here’s a new idea for treating advanced melanoma, the most serious kind of skin cancer. Genetically engineer white blood cells to better recognize and destroy cancer cells, then infuse these cells into patients.

An outcome in the cancer [experiment](JavaScript:top.Define('experiment')) described above is measured by a test for the presence of cells that trigger an immune response in the body and so may help fight cancer. Here is a simple random sample of 11 [subjects](JavaScript:top.Define('subject')): counts of active cells per 100,000 cells before and after infusion of the modified cells. The difference (after minus before) is the [response variable](JavaScript:top.Define('responsevariable')).

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Before | 14 | 0 | 1 | 0 | 0 | 0 | 0 | 20 | 1 | 6 | 0 |
| After | 41 | 7 | 1 | 215 | 20 | 700 | 13 | 530 | 35 | 92 | 108 |
| Difference(After-Before) | 27 | 7 | 0 | 215 | 20 | 700 | 13 | 510 | 34 | 86 | 108 |

(Remember this data is in Column 3, 4, 5 in R\_8\_Data)

* 1. **(1)** Make a probability plot (0.25pts) and a boxplot (0.25pts) of the differences and see if there are any major deviations from Normality in the differences (column 5). Is it reasonable to use the t procedures (0.25pts)? Why or why not(0.25 pts)? (Copy and paste the graphs that result with your answer. Make sure you have meaningful titles for your graph.)

***If the answer is no, do not do parts b.***

***A graph of a cell

AI-generated content may be incorrect.***

***A graph with a bar graph

AI-generated content may be incorrect.***

We observe firstly that there appears to be significant deviation from normality based on the qq-plot. As opposed to a linear pattern, the data appears to follow an exponential pattern on the plot. Furthermore, there are outliers present in the dataset as shown on the boxplot. Therefore neither of our conditions are met, and we will not continue with our t-procedures.

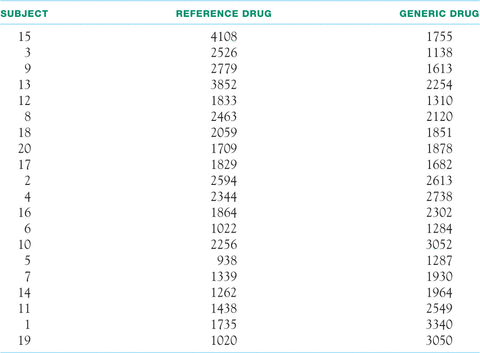
* 1. **(3pts, if needed to be done or 1.5 if you do and it didn’t need to be done) If** your conclusion in part (a) is **yes**, do the data give convincing evidence that the count of active cells is higher after treatment at a 0.05 significance level.
     1. (0.5pts) State the null and alternative hypotheses.
     2. (0.5pts) State the significance level for this problem.
     3. (0.5pts) State the test statistic. Use R to compute the test statistic and *P*-value. (Copy and paste the results. Also, use this result to answer part (iv)).
     4. (0.5pts) State the *P*-value.
     5. (0.5pts) State whether you reject or do not reject the null hypothesis.
     6. (0.5pts) State your conclusion in context of the problem.

**N/A**

1. The data set in Columns 7, 8, 9, and 10 labeled Subj, Reference, Generic, Difference\_3.

Makers of generic drugs must show that they do not differ significantly from the “reference” drugs that they imitate. One aspect in which drugs might differ is their extent of absorption in the blood. [Table 17.6](JavaScript:top.OpenSupp('table','17',6)) gives data taken from 20 healthy nonsmoking male [subjects](JavaScript:top.Define('subject')) for one pair of drugs. This is a [matched pairs](JavaScript:top.Define('matchedpairs')) design. Numbers 1 to 20 were assigned at random to the [subjects](JavaScript:top.Define('subject')). [Subjects](JavaScript:top.Define('subject')) 1 to 10 received the generic drug first, and [Subjects](JavaScript:top.Define('subject')) 11 to 20 received the reference drug first. In all cases, a washout period separated the two drugs so that the first had disappeared from the blood before the [subject](JavaScript:top.Define('subject')) took the second.

**TABLE 17.6** **Absorption extent for two versions of a drug**

[](javascript:top.OpenSupp('table',17,6))



(Remember this data is in column 7, 8, 9, and 10 in R\_8\_Data) You will note that subject 15 has been deleted as the input was incorrect.

* 1. (1) Make a probability plot (0.25pts) and a boxplot (0.25pts) of the differences and see if there are any major deviations from Normality in the differences. Is it reasonable to use the t procedures(0.25pts)? Why or why not? (0.25 pt) (Copy and paste the graphs that result with your answer. Make sure you have meaningful titles for your graph.)

A graph with dots on it

AI-generated content may be incorrect.

A graph of a bar chart

AI-generated content may be incorrect.

Fortunately, it appears that we can use our t-procedures here. Firstly the data appears to follow a linear pattern on the qq-plot, suggesting the distribution is approximately normal. Furthermore, there are no outliers evident from the boxplot. Thus our conditions are met and we can continue with t-procedures.

***If the answer is no, do not do part b.***

* 1. **(3 pts, if needed to be done or 1.5 if you do and it didn’t need to be done)If** your conclusion in part (a) is **yes**, do the data give convincing evidence that the drugs differ in absorption at a 0.05 significance level.
     1. State the null and alternative hypotheses.

Null Hypothesis: The mean difference in absorption of the drugs is equal to 0.

Alternative Hypothesis: The mean difference in absorption of the drugs is not equal to 0.

* + 1. State the significance level for this problem.

We are using a 0.05 significance level.

* + 1. State the test statistic to four decimal places. Use R to compute the test statistic and *P*-value. (Copy and paste the results found in the Session window after your answer. Also, use this result to answer part (iv)).

A white background with black text

AI-generated content may be incorrect.

The test statistic is approximately -0.6655.

* + 1. State the *P*-value to four decimal places.

The p-value is 0.5142.

* + 1. State whether you reject or do not reject the null hypothesis.

Since 0.5142 > 0.05, we do not reject the null hypothesis.

* + 1. State your conclusion in context of the problem.

We do not have convincing evidence to suggest that the mean difference of absorption of the drugs differs at a 0.05 significance level.