HW9: The Moving to Opportunity Experiment & Hypothesis Testing

Millions of low-income Americans live in high-poverty neighborhoods, which also tend to be racially segregated and sometimes have issues with community violence. While social scientists have long believed a lack of investment in these neighborhoods contributes to negative outcomes for the residents living in them, it is often difficult to establish a causal link between neighborhood conditions and individual outcomes. The Moving to Opportunity (MTO) demonstration was designed to test whether offering housing vouchers (that cover some or all of rental costs) to families living in public housing in high-poverty neighborhoods could lead to better experiences and outcomes by providing financial assistance to move to higher income neighborhoods.

Between 1994 and 1998 the U.S. Department of Housing and Urban Development enrolled 4,604 low-income households from public housing projects (governement owned housing, usually large apartment blocks) in Baltimore, Boston, Chicago, Los Angeles, and New York in MTO, *randomly assigning* enrolled families in each site to one of three groups: (1) The low-poverty voucher group received special MTO vouchers, which could only be used to cover rent for apartments/homes in census tracts with 1990 poverty rates below 10% and counseling to assist with relocation; (2) the traditional voucher group received regular section 8 vouchers, which they could use to cover rent for apartments/houses anywhere; and (3) the control group, who received no vouchers but continued to qualify for any project-based (government-owned) housing assistance they were entitled to receive. Today we will use the MTO data to learn if being given the opportunity to move to lower-poverty neighborhoods improved participants' economic well being. This exercise is based on the following article and the data is a subset of the data used for this article:

Ludwig, J., Duncan, G.J., Gennetian, L.A., Katz, L.F., Kessler, J.R.K., and Sanbonmatsu, L., 2012. "Neighborhood Effects on the Long-Term Well-Being of Low-Income Adults." *Science*, Vol. 337, Issue 6101, pp. 1505-1510.

NEW TO THIS HW:

Consider a policy setting where if the conclusion of this research study is that the Low Poverty Voucher program improves economic well being relative to the control and Section 8 programs (combined) then policy makers will *eliminate* the Section 8 program and replace it with an expanded Low Poverty Voucher Program. If the conclusion of this research study is that the Low Poverty Voucher Program seems to have a similar effect on economic well being as the Control and Section 8 programs (combined) then they will continue the Low Poverty Voucher Program at its current size and continue to collect more data to determine what the effects are in a larger study over a longer time period.

The file `mto2_HW9.csv` includes the variables listed below for adult participants in the voucher, control, and Section 8 groups. For this analysis we will consider the Low Poverty

Voucher program to be *the treatment* and the Control and Section 8 groups combined (usual housing assistance) to be *the Alternative*.

Name	Description
`group`	factor with 3 levels: `lpv` (low-poverty voucher), `sec8` (traditional section 8 voucher), and `control`

`econ_ss_zcore` | Standardized measure of economic self-sufficiency, centered around the control group mean and re-scaled such that the control group mean = 0 and its standard deviation = 1. Measure aggregates several measures of economic self-sufficiency or dependency (earnings, government transfers, employment, etc.)

`crime_vic` | Binary variable, `1` if a member of that household was the victim of a crime in the six months prior to being assigned to the MTO program, `0` otherwise based on self-report

The data we will use are not the original data as this dataset has been modified to protect participants' confidentiality, but the results of our analysis will be consistent with published data on the MTO demonstration.

```
library(tidyverse)
```

```
## ── Attaching core tidyverse packages ──────────────────────── tidyverse
2.0.0 ──
## ✓ dplyr     1.1.2     ✓ readr     2.1.4
## ✓ forcats   1.0.0     ✓ stringr   1.5.0
## ✓ ggplot2   3.4.3     ✓ tibble    3.2.1
## ✓ lubridate 1.9.2     ✓ tidyr     1.3.0
## ✓ purrr     1.0.2
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all
conflicts to become errors
```

```
library(tinytex)
```

```
mto2 <- read.csv("data/mto2_HW9.csv")
```

Question 1 [20 pts]

1a [3 points]

Create a dichotomous variable, *treat* that takes the value 1 if a household is in the Low Poverty Voucher group and takes the value 0 if a household is in either the Control group or the Section 8 Group (usual housing assistance). How many households in the data are in the treatment group? How many households in the data are in this combined alternative group?

1b [4 points]

What is the specific causal question the researchers wanted to answer? What is the missing counterfactual for one household in the Low Poverty Voucher treatment group?

1c [4 points]

State the Null and Alternative Hypotheses that the researchers want to test (use a two-sided alternative).

1d [3 points]

What is a Type 1 error for this Hypothesis test and what would the real-world consequences of a Type 1 error be given the actions policy makers would make as described above (see New To This HW section above)?

1e [3 points]

What is a Type 2 error for this Hypothesis test and what would the real-world consequences of a Type 2 error be given the actions policy makers would make as described above (see New To This HW section above)?

1f [3 points]

What alpha level would you recommend using for this hypothesis test and why?

Answer 1

Answer 1a

55.4 % are in the alternative group and 44.59 % are in the treatment group. 1808 is the number of members in the alternate group and 1455 are in the treatment (lpv) group.

```
econ_t <- mto2[mto2$treat==1,]
econ_a <- mto2[mto2$treat==0,]
mto2$treat <- ifelse(mto2$group=="lpv",1,0)
proportions(table(mto2$treat))

##
##         0         1
## 0.5540913 0.4459087

table(mto2$group)

##
## control     lpv    sec8
##    1136    1455     672

mean_samp <- mean(econ_t$econ_ss_zscore)
mean_samp

## [1] NaN
```

Answer 1b

SCQ: what is the impact on the economic well being for households that experience the Low poverty vouchers relative to recieving the sec 8 and control treatment ?

MCF : what the outcome on the economic well being would be if a household gets the treatment(lpv vouchers) instead receives sec8 and control alternative on the households that are from these american cities and all else remains the same.

Answer 1c

null and alternate hypothesis: if the conclusion of this research study is that the Low Poverty Voucher program improves economic well being relative to the control and Section 8 programs (combined) then policy makers will *eliminate* the Section 8 program and replace it with an expanded Low Poverty Voucher Program. If the conclusion of this research study is that the Low Poverty Voucher Program seems to have a similar effect on economic well being as the Control and Section 8 programs (combined) then they will continue the Low Poverty Voucher Program at its current size and continue to collect more data to determine what the effects are in a larger study

the null hypothesis : $\mu T - \mu A = 0$ μT mean of the treatment variable μA mean of the alternate variable (t indicates treatment groups = LPV receiving households) this means that under the null condition the mean of the economic well being has no change between the alternate and treatment groups (A indicates the SEC 8 and control groups combined)

THe alternate hypothesis: $\mu T - \mu A \mathrel{!}= 0$ (or $\mu T - \mu A > 0$)this means that under the alternate the change between the change in the mean economic well being is not zero between the treatment and alternate groups.

Answer 1d

type - 1 error: Type I error is the probability of falsely rejecting the null hypothesis when it is true; in this case, it is the probability of concluding that the treatment of LPV intervention changed (increased) the economic well being of households when infact it did not.A potential cost is that then policy makers will *eliminate* the Section 8 program and replace it with an expanded Low Poverty Voucher Program which wastes time and money.

Answer 1e

type -2 error: A Type II error is the probability of failing to reject the null hypothesis when the null hypothesis is false; in this case it is the probability of concluding that there was infact no change on the economic well being of households that received the control and sec 8 versus the treatment, even when there is change. a potential issue would be when the research study will continue the Low Poverty Voucher Program at its current size and continue to collect more data to determine what the effects are in a larger study. this is a failure to reject the null even when it is false, leading to money and time spent collecting more data.

Answer 1f

we can use the alpha level as 0.05 as there is no favourability.
it is preferable to have type 2 error than type 1 error this is because in type 1 error we scrap off the Sec8 and control and replace them with a program that is not as effective. this will effect the people that undergo the treatment and On the other hand if we have type 2 error all 3 programs will continue, the subjects that partake in the program will benefit with respect to their economic well being.

Question 2 [15 pts]

2a [5 points]

Make a histogram of the *econ_ss_zcore* outcome variable for households in the treatment group. Calculate summary statistics of this outcome variable for the treatment group. Briefly describe the distribution of economic self-sufficiency in the sample for treated households.

2b [5 points]

Consider the sampling distribution of the sample mean economic self-sufficiency variable for treated households over repeated realizations of the data from the data generating mechanism that created this data set. Does the CLT apply to this subgroup sample mean over hypothetical repeated sampling (repeated realizations of the DGM)? Explain why or why not. If yes, what are the shape, mean, and estimated standard error of this sampling distribution?

2c [3 points]

Calculate a 90% confidence interval for the mean economic self-sufficiency for those with a Low Poverty Voucher in the data generating mechanism (the unknown parameter of interest).

2d [2 points]

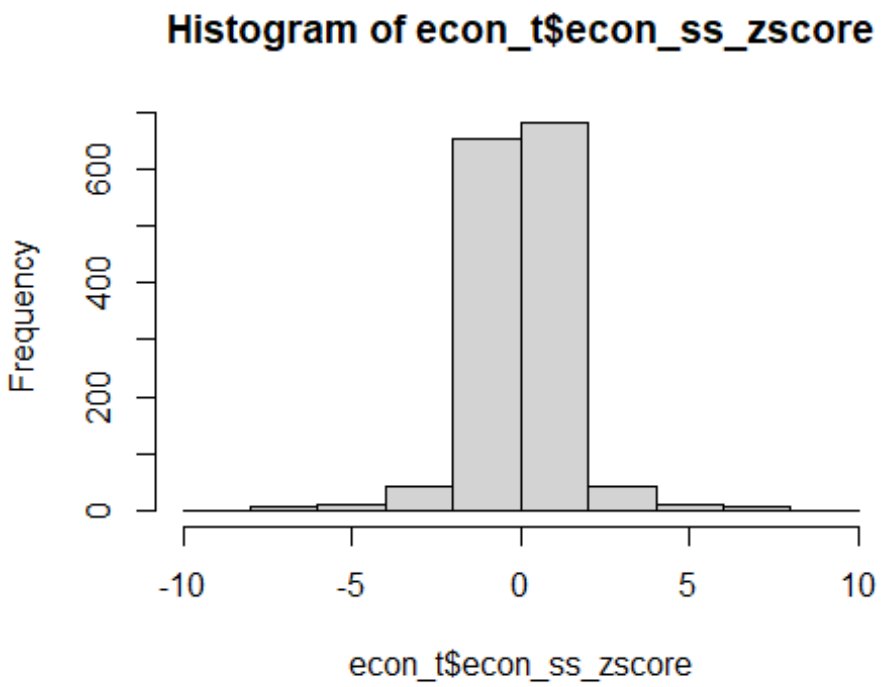
Give a statistical interpretation of this confidence interval.

Answer 2

Answer 2a

The distribution is symmetric, unimodal and has a bell shape.there seem to be no outliers. It tapers towards the edges and it dense only in the middle between the IQR . The economic sufficiency in the treatment group ranges from -9.09 to 8.39. The median is 0.005 and the mean is -0.022. The IQR is -0.14 to 0.13. It represents the economic z score of the treatment group.

```
econ_t <- mto2[mto2$treat==1,]
econ_a <- mto2[mto2$treat==0,]
hist(econ_t$econ_ss_zscore)
```




```r
summary(econ_t$econ_ss_zscore)
```

```
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
## -9.094809 -0.140514  0.005999 -0.022724  0.130949  8.391736
```

```r
sample_size_t <- nrow(econ_t)
sample_size_t
```

```
## [1] 1455
```

```r
sample_size_a <- nrow(econ_a)
sd(econ_t$econ_ss_zscore)
```

```
## [1] 1.165123
```

```r
mean(econ_t$econ_ss_zscore)
```

```
## [1] -0.02272391
```

```r
se_hat_t <- sd(econ_t$econ_ss_zscore)/sqrt(sample_size_t)
se_hat_t
```

```
## [1] 0.03054501
```

```r
se_hat_a <- sd(econ_a$econ_ss_zscore)/sqrt(sample_size_a)
se_hat_a
```

```
## [1] 0.03354768
```

Answer 2b

over repeated sampling, the CLT applies because the paramter of interest is a mean of the population. the mean of economic self sufficiency of the treatment group. the estimate standard error is 0.03. we expect this to follow CLT as there aren't extreme outliers and there isn't a heavy skew. Besides this the parameter of interest is the sample means of the economic well being score which is a MEAN and this follows and satisifes the CLT condition. Over repeated ssampling we expect this to take the shape of a normal distribution. It follows the CLT as the size is also greater than 500.

Answer 2c

this is the 90% confidence interval 90% CI: $\#X \pm 1.64 * SE$= (-0.07, 0.02)

```
ci_95_1 <- mean(econ_t$econ_ss_zscore) - 1.64 *se_hat_t
ci_95_1

## [1] -0.07281773

mean(econ_t$econ_ss_zscore)

## [1] -0.02272391

ci_95_2 <- mean(econ_t$econ_ss_zscore) + 1.64*se_hat_t
ci_95_2

## [1] 0.0273699
```

Answer 2d

over repeates The 90% confidence interval for the mean economic self-sufficiency for those with a Low Poverty Voucher in the data generating mechanism is (-0.07, 0.02). Over repeated realizations of data from this data generating mechanism, 90% of intervals constructed in this manner will contain the true average mean economic self sufficiency score when households are under treatment and 10% will not. ## Question 3 [12 pts]

3a [4 points]

In this question you will carry out a hypothesis test of whether or not the mean of the *econ_ss_zcore* outcome variable for treated households equals zero in the data generating mechanism. Use an alpha level of 0.10.

State the parameter of interest and the symbol you will use for it. State the estimator you will use for this parameter from the observed data (the observed data summary statistic). State the Null and Alternative Hypotheses.

3b [3 points]

Describe the sampling distribution of the sample summary statistic *under the Null Hypothesis* (give the shape, mean, and estimated standard error).

3c [1 points]

Calculate the number of standard errors the observed summary statistic is from the Null value of the parameter of interest.

3d [3 points]

Calculate the p-value for the hypothesis test. Interpret this p-value.

3e [1 point]

Give the conclusion of the hypothesis test.

Answer 3

Answer 3a

The parameter of interest is the mean of the economic well being score of the households under treatment of Low poverty vouchers. Let μA be the average mean economic well being score of households under treatment, in the data generating mechanism. We use the estimate of the mean of the treated group's economic well being score(econ_z score) across the households. The Null and Alternative Hypotheses are: H0 : μA = 0 Ha : μA != 0

Answer 3b

Under null hypothesis the mean economic score of households under the treatment condition is a t distribution, 1454 degrees of freedom and is a normal distribution. Estimated standard error: 0.03. Under the null hypothesis, the mean of this sample is 0.

```
mean(econ_t$econ_ss_zscore)

## [1] -0.02272391

summary(econ_t$econ_ss_zscore)

##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
## -9.094809 -0.140514  0.005999 -0.022724  0.130949  8.391736
```

Answer 3c

it is -0.74 estimated standard error s below the the estimated the sample mean of the treatment group

```
se_away_t<- (mean_samp - 0)/se_hat_t
se_away_t

## [1] NaN
```

Answer 3d

the p value is 0.45, the alpha value = 0.10 , therefore p>Alpha, and we would fail to reject the null hypothesis. we will not reject the null hypothesis. The p-value for this hypothesis

test is p =0.45 45% . The p-value is the probability that if the Null Hypothesis were true, that we would observe data with a mean this far from the Null Value (or any farther), just by chance.

```
2*(1 - pnorm(abs(-0.743)))
```

```
## [1] 0.4574817
```

```
2*(1 - pt(-0.743, sample_size_t - 1))
```

```
## [1] 1.542398
```

Answer 3e

Here that probability is 45 % much larger than our alpha level of 0.10, so we will fail to reject the Null Hypothesis (the economic mean well being level of households in the treatment is zero) in favor of the Alternative Hypothesis (that this mean is not zero).

Question 4 [9 pts]

4a [3 points]

Make a figure with side-by-side boxplots of the economic self-sufficiency outcome variable for the treatment (Low Poverty Voucher) group and the alternative group. Describe any differences you see between the two distributions.

4b [6 points]

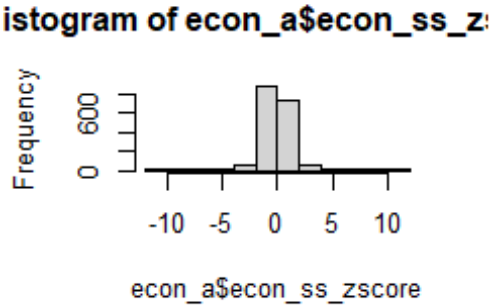
Not assuming the standard deviation of the outcome is the same in the treatment and alternative groups, calculate 95% and 99% confidence intervals for *the difference* in mean self-sufficiency if a household has a Low Poverty Voucher relative to an Alternative in the data generating mechanism. What is the estimated standard error of this sample difference in group means over repeated realizations of the data?

Answer 4

Answer 4a

the two are similarly distributed with a slight skew. The only notable difference is in their ranges. The treatment group ranges with economic well being values from -5 to +5 while the alternate ranges from -10 to 10 .

```
par(mfrow = c(2, 2))
boxplot(econ_t$econ_ss_zscore)
boxplot(econ_a$econ_ss_zscore)
hist(econ_a$econ_ss_zscore)
```



Answer 4b

the confidence interval for 95 % is : (-0.06801679, 0.10989347) the confidence interval for 99 % is : (-0.09, 0.13) the se is 0.045 for the difference of the means between alternate and treatment groups over repeated sampling.

```
se_hat_diff <- sqrt((se_hat_t)*(se_hat_t) + (se_hat_a)*(se_hat_a))
se_hat_diff

## [1] 0.04537008

mean_dif <- mean(econ_t$econ_ss_zscore) - mean(econ_a$econ_ss_zscore)
mean_dif

## [1] 0.02093995

se_hat_dif <- se_hat_t - se_hat_a
ci_diff<- c(mean_dif - qt(.975, sample_size_t + sample_size_a - 2) *
se_hat_diff,
mean_dif + qt(.975, sample_size_a + sample_size_a - 2) * se_hat_diff)
ci_diff_99 <- c(mean_dif - qt(.995, sample_size_t + sample_size_a - 2) *
se_hat_diff,
mean_dif + qt(.995, sample_size_a + sample_size_a - 2) * se_hat_diff)

ci_diff

## [1] -0.06801679  0.10989347
```

```
ci_diff_99
```

```
## [1] -0.09599407  0.13786729
```

```
se_hat_diff <- sqrt((se_hat_t)*(se_hat_t) + (se_hat_a)*(se_hat_a))
se_hat_diff
```

```
## [1] 0.04537008
```

Question 5 [12 pts]

5a [5 points]

Refer to your Null and Alternative Hypotheses from Q1c. Sketch the sampling distribution for the differences in sample means ((between the Low Poverty Voucher and Alternative groups) over repeated realizations of the data *under the Null Hypothesis*. What is the shape of this sampling distribution? Label the mean of this distribution and the points 1, 2 and 3 estimated standard errors away from the mean (using the estimated SE you calculated in Question 4B).

5b [4 points]

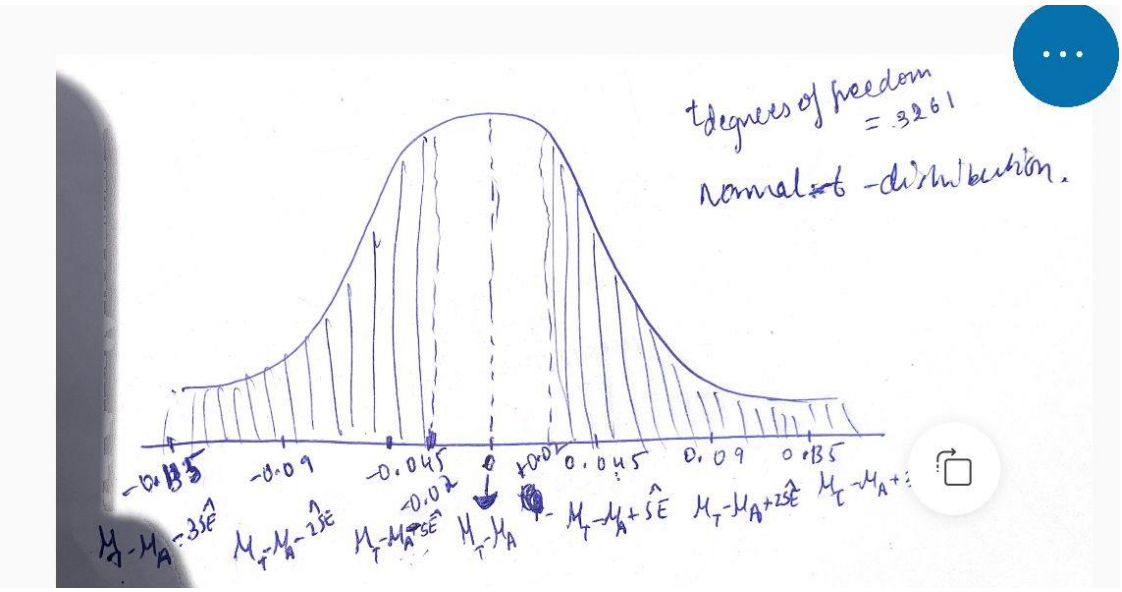
What is the mean difference we observe in this data? How many estimated standard errors away from the Null value is it? Put this point on your sketch of the sampling distribution and shade the area under the sampling distribution curve that corresponds to the p-value for this two-sided test.

5c [3 points]

Calculate and interpret the p-value for this hypothesis test. Based on the alpha level you recommended in Question 1 what is the conclusion of the hypothesis test?

Answer 5

Answer 5a



CAPTION

it is t-distribution that also is a normal distribution with 3261 degrees of freedom.

Answer 5b The mean difference is 0.02. It is at 0.46 SE away from the null value

```
se_away_2 = mean_dif/se_hat_diff
se_away_2
```

```
## [1] 0.4615366
```

```
mean_dif <- mean(econ_t$econ_ss_zscore) - mean(econ_a$econ_ss_zscore)
mean_dif
```

```
## [1] 0.02093995
```

Answer 5c

the p value is 0.644. It is the probability of 64.4 % that just by chance, we would observe data where the difference in means were this far from the null value (or any farther) if the Null Hypothesis were true.This is greater than the alpha we may select , we will fail to reject the null: that there is no difference in the means of the economic well being score between treatment and alternate groups, therefore there is in fact no true difference between the two treatment and alternative.

```
pvalue = 2*(1-pt(se_away_2, sample_size_a + sample_size_t - 2))
pvalue
```

```
## [1] 0.6444445
```

Question 6 [8 pts]

6a [3 points]

Use the *t.test* command to carry out the same hypothesis test you conducted in Question 5 (two-sided alternative hypothesis, equal standard deviations is not assumed). Compare the p-value results using the *t.test* command versus how you calculated the p-value in Question 5. Also compare the confidence interval given in the *t.test* output to the confidence interval with the same coverage level for the difference in means that you calculated in Question 4.

6b [3 points]

Use the *t.test* command to carry out the hypothesis test where the alternative is one sided instead of two-sided (which direction is consistent with the researchers hypothesis?) but otherwise is the same. Compare the results (the magnitude of the p-value and the degrees of freedom for the test) with the *t.test* results where the alternative hypothesis is two sided [the test you carried out in 6A]. What is the relationship between these two p-values?

6c [2 points]

Use the *t.test* command to carry out the 2-sided hypothesis test but now assuming the standard deviation is the same in the treatment and alternative groups. Compare the results (the magnitude of the p-value and the degrees of freedom for the test) with the *t.test* results where equal standard deviations is not assumed [the test you carried out in 6A].

Answer 6

Answer 6a

The p value in q5 is the same as compared to the p-value calcultated in 6. (-0.1098967 0.0680168) is the CI for t-test two sided in q6 and the one in q4 is (-0.06801679 0.10989347) same width and length as the CI we calculated in the fourth question for 95 percent confidence interval. although the signs have been reversed due to the internal coding of R programming as the formula to calculate is different but the length of the interval is the same. similarly, the Ci for the 99 % confidence interval is the same between the 4 th and the 6 th question -> (-0.09, 0.13).

```
t.test( econ_ss_zscore ~ treat,
data = mto2,
alternative = 'two.sided',
var.equal = FALSE,
conf.level = 0.95)

##
##  Welch Two Sample t-test
##
## data:  econ_ss_zscore by treat
## t = -0.46154, df = 3260.3, p-value = 0.6444
## alternative hypothesis: true difference in means between group 0 and group
1 is not equal to 0
```

```
## 95 percent confidence interval:
##  -0.1098967  0.0680168
## sample estimates:
## mean in group 0 mean in group 1
##     -0.04366387     -0.02272391
```

```
t.test( econ_ss_zscore ~ treat,
data = mto2,
alternative = 'two.sided',
var.equal = FALSE,
conf.level = 0.99)
```

```
##
##  Welch Two Sample t-test
##
## data:  econ_ss_zscore by treat
## t = -0.46154, df = 3260.3, p-value = 0.6444
## alternative hypothesis: true difference in means between group 0 and group
1 is not equal to 0
## 99 percent confidence interval:
##  -0.13787399  0.09599409
## sample estimates:
## mean in group 0 mean in group 1
##     -0.04366387     -0.02272391
```

Answer 6b

The p value is 0.322 in 6b and the p value is 0.64 for 6a and df for 6b is 3260. The p value has changed significantly. the p-value is less than what we calculated in 6a. Therefore we have evidence to fail to reject the null. The p value is half the value we calculated in 6a and this is because we considered the greater than side i.e the +ve side of the null hypothesis. The degrees of freedom is 3260.

```
t.test( econ_ss_zscore ~ treat,
data = mto2,
alternative = 'less',
var.equal = FALSE,
conf.level = 0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  econ_ss_zscore by treat
## t = -0.46154, df = 3260.3, p-value = 0.3222
## alternative hypothesis: true difference in means between group 0 and group
1 is less than 0
## 95 percent confidence interval:
##       -Inf 0.0537084
## sample estimates:
## mean in group 0 mean in group 1
##     -0.04366387     -0.02272391
```

Answer 6c

p value is similar 0.65 and has slightly increased compared to the p value 0.644 we calculated in 6a. The degrees of freedom has increased from 3260 from 6a to 3261 in 6c.

```
t.test( econ_ss_zscore ~ treat,
data = mto2,
alternative = 'two.sided',
var.equal = TRUE,
conf.level = 0.95)

##
##  Two Sample t-test
##
## data:  econ_ss_zscore by treat
## t = -0.45167, df = 3261, p-value = 0.6515
## alternative hypothesis: true difference in means between group 0 and group
1 is not equal to 0
## 95 percent confidence interval:
##  -0.11183989  0.06995999
## sample estimates:
## mean in group 0 mean in group 1
##     -0.04366387     -0.02272391
```