

Victoria Haley
Homework 5

Exercise 6

R code:

```
t.test(PlantGrowth$weight[PlantGrowth$group=='ctrl'],PlantGrowth$weight[PlantGrowth$group=='trt1'])
```

Output:

t-value: 1.1913 df: 16.524 p-value: 0.2504
upper bound: 1.0295162 lower bound: -0.2875162

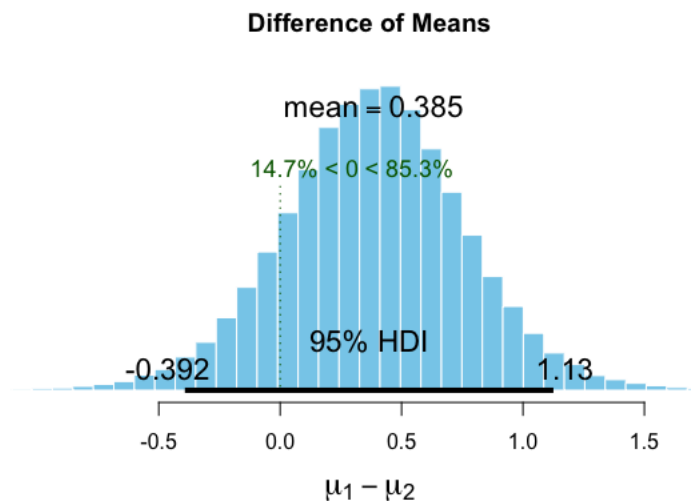
Interpretation:

Assuming an alpha level of 0.05, we should fail to reject the null hypothesis. Seeing that the p-value is greater than the chosen alpha level, we do not have evidence to rule out the null hypothesis.

Exercise 7

R code:

```
library(BEST)
PlantGrowthBest <-
BESTmcmc(PlantGrowth$weight[PlantGrowth$group=='ctrl'],PlantGrowth$weight[PlantGrowth$group=='trt1'])
plot(PlantGrowthBest)
output:
```



HDI boundary values: -0.392 and 1.13.

Interpretation:

There is a 95% chance that the mean difference falls between -0.392 and 1.13. The HDI is the range where the difference of means is 95% likely to be between. The HDI of the MCMC of the 'ctrl' and 'trt1' group tells us that 85.3% of the steps were positive, while 14.7% were negative, and had a mean of 0.385. This further shows that the null hypothesis fails to be rejected as a mean difference of 0 is a possible outcome.

Exercise 8

The HDI of the test shows that the mean difference in weight between plants in 'ctrl' and 'trt1' groups is 0.385 and favors the weights of the 'trt1' group. The Confidence interval from above indicates that the difference of mean weights is somewhere between -0.2875162 and 1.0295162, however this does not guarantee that the population parameter is contained within that range. The null hypothesis fails to be rejected in this test, which is fair because there is a chance that there is zero difference between the 'ctrl' and 'trt1' groups.

Exercise 9

R code:

```
t.test(PlantGrowth$weight[PlantGrowth$group=='ctrl'],PlantGrowth$weight[PlantGrowth$group=='trt2'])
```

Output:

```
t-value: -2.134      df: 16.786      p-value: 0.0479
upper bound: -0.98287213  lower bound: -0.00512787
```

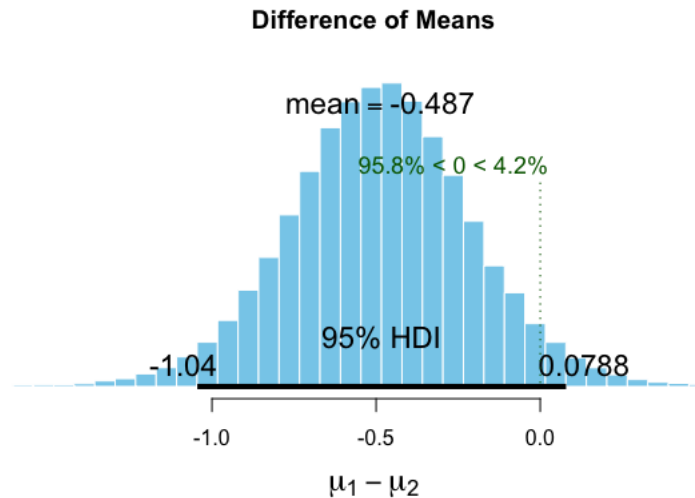
Interpretation:

Assuming an alpha level of 0.05, the null hypothesis should be rejected here. Seeing that the p-value is less than the chosen alpha level, there is evidence to support that the null hypothesis is not likely.

R code:

```
PlantGrowthBest2 <-
BESTmcmc(PlantGrowth$weight[PlantGrowth$group=='ctrl'],PlantGrowth$weight[PlantGrowth$group=='trt2'])
plot(PlantGrowthBest2)
```

output:



HDI boundary values: -1.04 and 0.0788.

Interpretation:

There is a 95% chance that the mean difference falls between -1.04 and 0.0788. The HDI of the MCMC of the 'ctrl' and 'trt2' group tells us that 95.8% of the steps were negative, while 4.2% were positive, and had a mean of -0.487, meaning that. While it is possible for there to be zero difference in the means, it is not very likely.

The HDI of the test shows that the mean difference in weight between plants in 'ctrl' and 'trt2' groups is 0-0.487 and favors the weights of the 'ctrl' group. The Confidence interval from above indicates that the difference of mean weights is somewhere between -0.00512787 and -0.98287213, however this does not guarantee that the population parameter is contained within that range. The null hypothesis can be rejected in this test, meaning that we can say that this is evidence in favor of an alternative hypothesis.

Exercise 10

R code: `t.test(rnorm(100000, mean=17.1,sd=3.8),rnorm(100000,mean=17.2,sd=3.8))`

Output:

Welch Two Sample t-test

data: `rnorm(1e+05, mean = 17.1, sd = 3.8)` and `rnorm(1e+05, mean = 17.2, sd = 3.8)`

t = -4.9856, df = 2e+05, p-value = 6.184e-07

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-0.11808654 -0.05144026

sample estimates:
mean of x mean of y
17.11366 17.19842

Interpretation:

The difference in the means from above ($17.11366 - 17.19842$) is -0.08476 which is not the same as the difference of the mean in the data above ($17.1 - 17.2$), which is only -0.1 . Although -0.08476 is not very far from -0.1 , the difference could mean failing or succeeding when trying to determine statistical significance. Using the NHST on large data sets could affect the meaning of a piece of research as any observation could be trivial, but we are unable to determine anything about what is going on.