

```
#####
#
# Author: DeAndre Thomas
# Professor: Dr. Anastasopoulos
# Purpose: Homework 5
# Class: IST 772
# Date: 05/13/2022
#
#####

# Question 6

View(PlantGrowth)

t.test(PlantGrowth$weight[PlantGrowth$group=="ctrl"]
      , PlantGrowth$weight[PlantGrowth$group=="trt1"])
# Welch Two Sample t-test

# data: PlantGrowth$weight[PlantGrowth$group == "ctrl"]
# and PlantGrowth$weight[PlantGrowth$group == "trt1"]
# t = 1.1913, df = 16.524, p-value = 0.2504
# alternative hypothesis: true difference in means is not equal to 0
# 95 percent confidence interval:
# -0.2875162 1.0295162
# sample estimates:
# mean of x mean of y
# 5.032 4.661

# t = 1.1913
# degrees of freedom (df) = 16.524
# p-value = 0.2504
# Assuming an alpha threshold of .05, we should fail to reject the null
# hypothesis.
# Lower Bound: -.0288 Upper Bound: 1.03

#####

# Question 7

install.packages("BEST")
library(BEST)

best1<- BESTmcmc(PlantGrowth$weight[PlantGrowth$group=="ctrl"]
               , PlantGrowth$weight[PlantGrowth$group=="trt1"])
best1
# MCMC fit results for BEST analysis:
# 100002 simulations saved.
# mean sd median HDIlo HDIup Rhat n.eff
# mu1 5.0259 0.2243 5.0253 4.5759 5.469 1.000 57186
# mu2 4.6418 0.3070 4.6387 4.0291 5.255 1.000 56802
# nu 34.5830 29.7692 26.0197 1.3176 93.929 1.002 20918
# sigma1 0.6603 0.2000 0.6228 0.3413 1.061 1.000 28547
# sigma2 0.8966 0.2774 0.8449 0.4629 1.461 1.000 27035

# 'HDIlo' and 'HDIup' are the limits of a 95% HDI credible interval.
# 'Rhat' is the potential scale reduction factor (at convergence, Rhat=1).
```

```

# 'n.eff' is a crude measure of effective sample size.

plot(best1)

# The boundary values that BESTmcmc() calculated for the HDI are -0.378 and
# 1.13 when comparing the control and treatment 1 groups. The HDI
# (or High Density Interval), similar to a confidence interval, give us
# the tools to understand the range of difference in the population mean.
# However, the HDI is more detailed. For example, in this specific program,
# the HDI tells us that there is a 95% chance that the population mean
# difference between the control and treatment 1 groups falls between -0.378
# and 1.13, while there is a 14.7% chance that the value is less than 0 and
# a 85.3% chance that the value is greater than 0.

#####

# Question 8

# The null hypothesis test tells us that with 95% confidence, we should assume
# that the growth of the control group should be equal to that of treatment 2
# group. The confidence intervals tell us that the population mean of both
# groups should be between -0.0288 and 1.03. The HDI tells us that we should
# highly consider (more than the CI), with 95% confidence, that
# the population mean will be between -0.378 and 1.13, while 14.7% of
# the population mean values will be less than 0 and the remaining
# 85.3% will be greater than 0.

# With this being noted, it is safe to assume, with 95% confidence, that the
# growth of plants in treatment 1 will have an 85.3% chance of growing more
# than those in the control group, with a population mean of 0.384.

#####

# Question 9

t.test(PlantGrowth$weight[PlantGrowth$group=="ctrl"]
      , PlantGrowth$weight[PlantGrowth$group=="trt2"])

# t = -2.134
# degrees of freedom (df) = 16.786
# p-value = 0.0479
# Assuming an alpha threshold of .05, we should reject the null hypothesis.
# Lower Bound: -0.9829      Upper Bound: -0.0051

# Welch Two Sample t-test

# data: PlantGrowth$weight[PlantGrowth$group == "ctrl"]
# and PlantGrowth$weight[PlantGrowth$group == "trt2"]
# t = -2.134, df = 16.786, p-value = 0.0479
# alternative hypothesis: true difference in means is not equal to 0
# 95 percent confidence interval:
# -0.98287213 -0.00512787
# sample estimates:
# mean of x mean of y
# 5.032      5.526

```

```

best2<- BESTmcmc(PlantGrowth$weight[PlantGrowth$group=="ctrl"]
                , PlantGrowth$weight[PlantGrowth$group=="trt2"])
best2
# MCMC fit results for BEST analysis:
# 100002 simulations saved.
# mean      sd  median HDIlo  HDIup  Rhat n.eff
# mu1      5.0264 0.2254 5.0257 4.5774 5.4793 1.000 57265
# mu2      5.5142 0.1710 5.5123 5.1798 5.8602 1.000 53983
# nu      34.6532 29.8357 25.9874 1.0455 93.8817 1.000 20182
# sigma1   0.6606 0.2027 0.6227 0.3436 1.0615 1.001 24704
# sigma2   0.5016 0.1571 0.4726 0.2561 0.8144 1.001 24223

# 'HDIlo' and 'HDIup' are the limits of a 95% HDI credible interval.
# 'Rhat' is the potential scale reduction factor (at convergence, Rhat=1).
# 'n.eff' is a crude measure of effective sample size.

plot(best2)

# The boundary values that BESTmcmc() calculated for the HDI are -1.05 and
# 0.0657 when comparing the control and treatment 1 groups. The HDI
# (or High Density Interval), similar to a confidence interval, give us
# the tools to understand the range of difference in the population mean.
# However, the HDI is more detailed. For example, in this specific program,
# the HDI tells us that there is a 95% chance that that the population mean
# difference between the control and treatment 2 groups falls between -1.05
# and 0.0657, while there is a 95.8% chance that the value is less than 0 and
# a 4.2% chance that the value is greater than 0.

#####

# Question 10

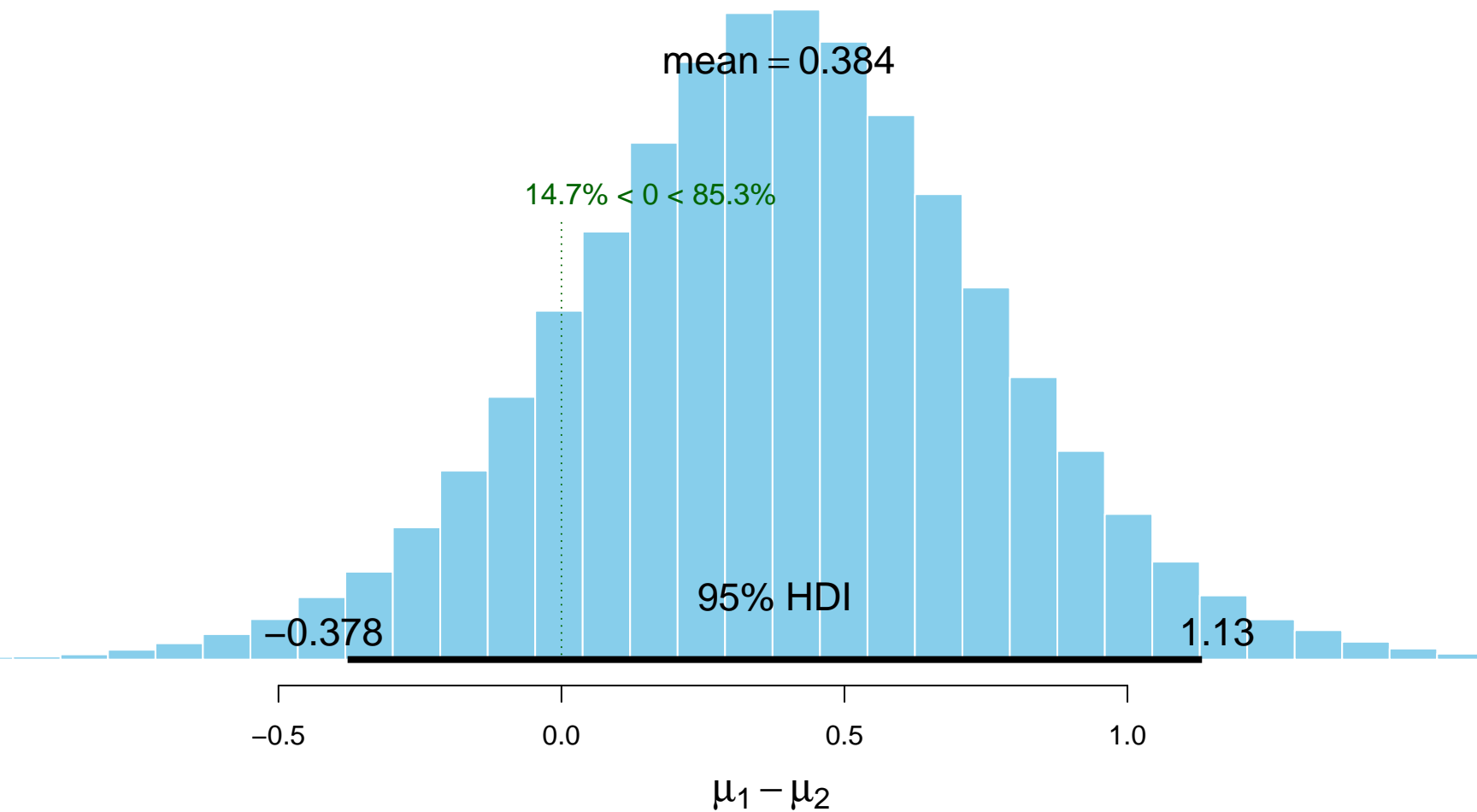
t.test(rnorm(100000,mean=17.1,sd=3.8),rnorm(100000,mean=17.2,sd=3.8))
# Welch Two Sample t-test

# data:  rnorm(1e+05, mean = 17.1, sd = 3.8) and rnorm(1e+05, mean = 17.2, sd
# = 3.8)
# t = -5.1199, df = 199993, p-value = 3.06e-07
# alternative hypothesis: true difference in means is not equal to 0
# 95 percent confidence interval:
# -0.12055078 -0.05380471
# sample estimates:
# mean of x mean of y
# 17.09324 17.18042

# By running this command, we can see that the p-value is extremely low,
# meaning that the null hypothesis would be rejected. Nonetheless, one
# could assume that being that the means of the two sets being compared
# are very similar, we would fail to reject the null hypothesis. With this
# being the case, one could imply that that using the NHST on very large
# data sets is an unreliable method of comparing sets of data.

```


Difference of Means



Difference of Means

