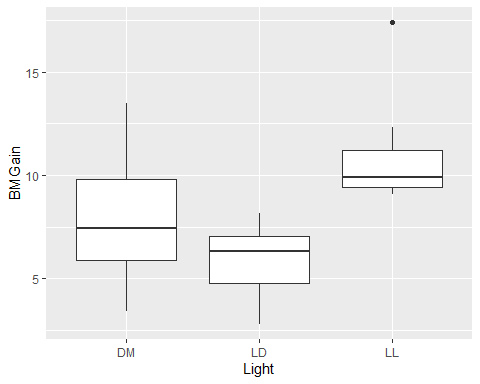
Homework 2

Xinyi Lin

9/22/2019

## Question 1



According to the plot, we can find that the average of weight gain in bright light group is higher than that in dim light group and the dark light group. The difference of weight gain in dim light group is the highest.

## Question 2

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Light | BMGain | Corticosterone | DayPct | Consumption | GlucoseInt | GTT15 | GTT120 | Activity |
| LD | 5.02 | 87.838 | 31.063 | 3.791 | No | 228.448 | 134.483 | 1437 |
| LD | 6.67 | 191.220 | 41.408 | 3.923 | No | 231.183 | 220.430 | 2541 |
| LD | 8.17 | 67.700 | 47.573 | 4.489 | No | 226.563 | 141.406 | 346 |
| LD | 2.79 | 41.017 | 34.947 | 4.161 | No | 323.077 | 199.038 | 5837 |
| LD | 8.13 | 21.817 | 41.940 | 4.416 | No | 500.000 | 190.361 | 877 |
| LD | 6.34 | 23.403 | 40.500 | 4.890 | No | 280.000 | 118.333 | 1649 |
| LD | 6.32 | 70.470 | 28.950 | 4.946 | No | 299.174 | 153.719 | 728 |
| LD | 3.97 | 56.718 | 21.846 | 4.004 | No | 461.250 | 230.000 | 6048 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Light | BMGain | Corticosterone | DayPct | Consumption | GlucoseInt | GTT15 | GTT120 | Activity |
| LL | 9.89 | 42.132 | 71.552 | 3.387 | Yes | 378.704 | 328.704 | 5752 |
| LL | 9.58 | 48.238 | 61.453 | 3.451 | No | 379.091 | 227.273 | 1256 |
| LL | 11.20 | 92.191 | 85.978 | 3.501 | Yes | 366.129 | 383.871 | 244 |
| LL | 9.05 | 51.999 | 64.827 | 4.240 | No | 392.373 | 250.000 | 931 |
| LL | 12.33 | 12.252 | 81.600 | 3.479 | Yes | 466.346 | 470.192 | 3582 |
| LL | 9.39 | 3.000 | 87.257 | 5.940 | Yes | 259.615 | 413.462 | 2657 |
| LL | 10.88 | 132.400 | 70.441 | 4.586 | No | 348.780 | 126.016 | 153 |
| LL | 9.37 | 8.615 | 84.415 | 4.873 | Yes | 335.652 | 286.957 | 4482 |
| LL | 17.40 | 66.679 | 81.636 | 7.177 | Yes | 435.644 | 405.941 | 6702 |

Records of subset datasets of dark light group and bright light group are shown above. There are 8 records in dark light group data and 9 records in bright light group data.

### Question 3

The observed outcomes: , where stands for individuals and stands for treatment.

The treatment assignment vector: , where . stands for being assigned to darkness and stands for being assigned to bright light.

The covariates:

### Question 4

### Question 5

## [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12] [,13]  
## [1,] 0 0 0 0 0 0 0 0 1 1 1 1 1  
## [2,] 0 0 0 0 0 0 0 1 0 1 1 1 1  
## [3,] 0 0 0 0 0 0 0 1 1 0 1 1 1  
## [4,] 0 0 0 0 0 0 0 1 1 1 0 1 1  
## [5,] 0 0 0 0 0 0 0 1 1 1 1 0 1  
## [6,] 0 0 0 0 0 0 0 1 1 1 1 1 0  
## [,14] [,15] [,16] [,17]  
## [1,] 1 1 1 1  
## [2,] 1 1 1 1  
## [3,] 1 1 1 1  
## [4,] 1 1 1 1  
## [5,] 1 1 1 1  
## [6,] 1 1 1 1

The first 6 possibilities are shown above. There are 24310 different possibilities are there for A.

### Question 6

The sharp null hypothesis is there is no treatment effect:

which means the vector of observed outcomes Y does not change with different .

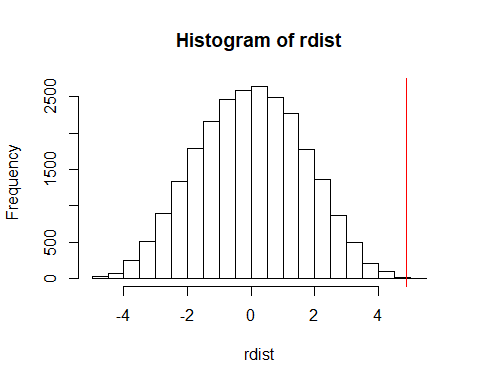
Calculating by R, we can get the test statistic under the first possibilities for A is 1.552.

### Quetion 7

## [1] 1.551527778 -0.700972222 0.290694444 -0.006805556 -0.358611111  
## [6] 0.713333333 -0.002083333 0.718055556 -1.898055556 1.272916667  
## [11] 0.569305556 1.560972222 1.263472222 0.911666667 1.983611111  
## [16] 1.268194444 1.988333333 -0.627777778 2.543194444 -0.691527778

The first 20 values of possible are shown above.

### Question 8



The plot of T’s distribution and the observed test statistic are shown above.

### Question 9

The exact p-value based on this distribution is 4.113533510^{-5}.

### Question 10

According to the plot and caculated p-value, we can find that the p-value is pretty small, which means the sharp null hypothesis needs to be rejected and there is treatment effect on outcomes.

**Appendix**

knitr::opts\_chunk$set(echo = FALSE)  
library(ggplot2)  
library(tidyverse)  
library(perm)  
library(ri)  
light = read.csv("./light.csv")  
# Quetion 1  
light %>%   
 ggplot(aes(x = Light, y = BMGain)) +  
 geom\_boxplot()  
# Question 2  
subset\_LD = light %>%   
 filter(Light == "LD")  
subset\_LL = light %>%   
 filter(Light == "LL")  
subset\_LD %>% knitr::kable()  
subset\_LL %>% knitr::kable()  
# Question 4  
T\_obs = mean(subset\_LL$BMGain) - mean(subset\_LD$BMGain)  
# Question 5  
Amatrix = chooseMatrix(17, 9)   
Amatrix %>%   
 head()  
# Question 6  
subset = light %>%   
 filter(Light %in% c("LD", "LL"))  
Y = subset$BMGain  
A\_tilde = Amatrix[1,]  
test\_sta = round(mean(Y[A\_tilde==1]) - mean(Y[A\_tilde==0]), 3)  
# Question 7  
rdist = rep(NA, times = nrow(Amatrix))  
for(i in 1:nrow(Amatrix)){  
 A\_tilde = Amatrix[i,]  
 rdist[i] = mean(Y[A\_tilde==1]) - mean(Y[A\_tilde==0])  
}  
rdist[c(1:20)]  
# Question 8  
pval = mean(rdist>=T\_obs)  
quant = quantile(rdist, probs = 1-pval)  
hist(rdist)  
abline(v = quant, col = "red")