MET CS 555

Assignment 2

Hao Wu

09/29.2021

(1)

图表, 直方图

描述已自动生成图表, 直方图

描述已自动生成

Participants

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Min.** | **X1st.Qu.** | **Median** | **Mean** | **X3rd.Qu.** | **Max.** |
| **1** | 210.99 | 298.38 | 424.94 | 410.0796 | 456.3 | 635.21 |

Nonparticipants

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Min.** | **X1st.Qu.** | **Median** | **Mean** | **X3rd.Qu.** | **Max.** |
| **1** | 139.69 | 296.3925 | 374.74 | 374.071818181818 | 445.5575 | 688.77 |

The histogram show us the data is normal distribution; however, the histogram of non\_particitants is right skew distribution. In the histogram of participants is left skew distribution.

(2)

The mean calorie consumption for those who participated is 425

H0:μ =425

H1:μ ≠ 425.

Degree of freedom = N – P

Because we only estimating one parameter(mean), so that p is 1.

Degree of freedom = 25 -1 = 24

The critical value is 1.710882

Test statistics is 0.61394 and the p-value is 0.545, so that there is not statistic significant, we can’t reject null hypothesis.

Conclusion:

the mean calorie consumption for those who participated in the meal preparation no differ from 425

(3)

90 percent confident interval is (368.5004, 451.6588)

(4)

H0:μparticipants = μnonparticipants

H1:μparticipants ≠ μnonparticipants

Degree of freedom = (25+22)-2 = 45

The critical value is 1.679427

Test statistics is 0.9636 and p-value is 0.3406, so that there is not statistic significant, we can’t reject null hypothesis.

Conclusion:

The mean of whose participated in the meal preparation no differ from whose no participated in the meal preparation.

(5)

The assumptions of the test used in (4) not met, because the data number from non-participated different from participated.

Code:

df1 <- read.csv("/Users/haowu/Desktop/Boston University Graduate Study/CS 555/assignment 2/Calorie.csv", header = FALSE)

typeof(df1)

names(df1) <- c("participants","non\_participants")

df2 <- df1[,c(1)]

df3 <- df1[,c(2)]

hist(df2, main="histogram of the participants", xlab = "Children's calories", ylab = "Frequency", border = "black", freq = TRUE)

hist(df3, main="histogram of the non\_participants", xlab = "Children's calories", ylab = "Frequency", border = "black", freq = TRUE)

df4 <- as.list(summary(df2))

write.csv(df4, file ="Statistic Summary of Participants.csv")

df5 <- as.list(summary(df3))

write.csv(df5, file ="Statistic Summary of non\_participants.csv")

qt(0.05,24)

t.test(df2,mu=425,alternative = "two.sided",conf.level = 0.95)

t.test(df2,conf.level = 0.90)

qt(0.05,45)

t.test(df2,df3)

图表

描述已自动生成